

U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL METEOROLOGICAL CENTER

OFFICE NOTE 203

Regional Multivariate Optimum Interpolation Analysis

Kenneth H. Bergman
Doris S. Gordon
Development Division

JUNE 1979

This is an unreviewed manuscript, primarily
intended for informal exchange of information
among NMC staff members.

ABSTRACT

The design of a regional multivariate optimum interpolation analysis scheme is presented. The way in which the error correlations of the guess fields and observations are analytically modeled is described. The moisture representation is obtained by univariate analysis of relative humidity for the three lowest sigma layers of the LFM model. Data selection and error-checking procedures are discussed. Results for one case suggest that the new scheme is capable of producing analyses and forecasts of comparable skill to those produced by the currently operational LFM analysis, with perhaps an improvement in precipitation forecasts.

1. Introduction

A multivariate analysis scheme based on the optimum interpolation method developed by Gandin (1963) and others has been designed and used for experimental LFM analyses at NMC. Many features of this scheme are very similar to, if not identical with, the multivariate analysis scheme used for global data assimilation and documented in NMC Office Note 200 by Bergman (1979). This office note will concentrate on the ways in which the regional scheme differs from the global scheme.

The theory of multivariate optimum interpolation (MOI) analysis in its most comprehensive form is presented in Gandin and Kagan (1974). MOI analysis schemes for height and wind fields have been documented by Schlatter (1975) and Rutherford (1973, 1976). These analysis schemes served as models when the NMC global MOI scheme was designed. The global scheme analyzes the temperature and the two horizontal wind components multivariately and the specific humidity univariately. These variables were chosen for analysis because they are the variables actually predicted at each time step by the nine-layer global prediction model which serves as data-assimilator (McPherson et al., 1977). The global analysis is performed on a longitude-latitude grid, and the variables are analyzed at the midpoints of the sigma-layers of the nine-layer global model. A 6-hour global prediction is used as the "guess" field for the analysis, and the fields of observed-minus-forecast differences ("residuals") are analyzed by the scheme, with a geostrophic thermal wind relationship between temperature residuals used in the wind field analysis and vice versa. The analysis is three-dimensional in that observations in a volume about the grid point on a sigma surface may be used in the analysis, rather than just observations appearing on some two-dimensional surface.

This three-dimensional concept of analysis is also used in the regional MOI scheme. But the regional scheme differs from the global one in the following significant ways:

1. Geopotential height, temperature, and the two horizontal wind components are analyzed multivariately, but with some restrictions on use of essentially redundant data (e.g., rawinsonde heights vs. temperatures). Relative humidity is analyzed univariately as the moisture variable.
2. The analysis is performed on a Cartesian grid superimposed on a polar stereographic projection. The grid is currently the LFM-I grid, but orientation and spacing may be readily changed.
3. The geopotential height, temperature, and wind components are analyzed at the mandatory isobaric levels, but the relative humidity is analyzed to the midpoints of the three lowest sigma layers of the LFM prediction model.

4. The guess fields for height, temperature, and relative humidity are provided by the hemispheric prediction model (12-hr forecasts). The guess wind field is generated by blending one-half the winds from the 12-hour hemispheric prediction with one-half the gradient winds computed from the predicted heights. The gradient wind formula used is that of McDonnell (1973).

5. The height analysis and wind residuals, and vice versa, are related through the geostrophic wind equation. Heights and temperatures are related by the hydrostatic equation. The thermal wind equation relates temperatures and winds as before.

Observational error characteristics are specified and used in each analysis scheme, but here again there are differences. In the global scheme, the actual magnitude of rms observational errors is specified. When used in the analysis, the rms observational error is compared with the local estimated rms error of the guess field. The latter error estimate evolves in time through the series of analyses and 6-hr forecasts produced by the cycling of the system. The ratio of the observational error to the guess error is the quantity that actually appears in the analysis equations. In the regional scheme, only this ratio of errors is specified, since specific information about errors in the guess is not available, and since local estimates of guess error cannot be generated in this noncycling system.

The basic equation for multivariate interpolation of n_k observations of each meteorological variable k to obtain the grid point correction to the r^{th} meteorological variable, f_{gr} , is given by

$$f_{gr} \approx \sum_{k=1}^m \sum_{i=1}^{n_k} a_{ik} \hat{f}_{ik}, \quad (1.1)$$

where

$$\hat{f}_{ik} = f_{ik} + e_{ik}, \quad (1.2)$$

and where a_{ik} is the combined weight/scaling factor which the i^{th} observation of meteorological variable type k receives in the interpolation, \hat{f}_{ik} is the observed residual (observed-minus-forecast value), f_{ik} is the true correction which should be applied at the observing point, and e_{ik} is the rms estimate of the observational error. Since four variables are analyzed multivariately by the regional scheme, subscript k assumes the values z , t , u , and v .

The weight/scaling factors a_{ik} are determined by solution of the set of equations:

$$\sum_{\ell=1}^m \sum_{j=1}^{n_{\ell}} (\rho_{ij}^{k\ell} + \eta_{ij}^{k\ell} \epsilon_{ik} \epsilon_{j\ell}) a_{j\ell} = \rho_{ig}^{kr}, \quad (1.3)$$

$$k = 1, 2, \dots, m; \quad i = 1, 2, \dots, n_k.$$

The theoretical estimate of mean-square analysis error of the r^{th} variable at grid point g is then given by

$$E_{gr}^2 / \overline{(f_{gr}^2)} = 1 - \sum_{k=1}^m \sum_{i=1}^{n_k} a'_{ik} \rho_{ig}^{kr} . \quad (1.4)$$

In the above equations,

$$\rho_{ij}^{kl} \equiv \overline{(f_{ik} f_{jl})} / \overline{(f_{ik}^2 f_{jl}^2)}^{1/2} , \quad (1.5)$$

$$\eta_{ij}^{kl} \equiv \overline{(e_{ik} e_{jl})} / \overline{(e_{ik}^2 e_{jl}^2)}^{1/2} , \quad (1.6)$$

$$\epsilon_{ik} \equiv \overline{(e_{ik}^2 / f_{ik}^2)}^{1/2} , \quad (1.7)$$

and

$$a'_{ik} = \overline{(f_{ik}^2 / f_{gr}^2)}^{1/2} a_{ik} . \quad (1.8)$$

Equations (1.3) and (1.4) are the same as (2.12) and (2.13) of Office Note 200.

The above equations are the starting point for the regional MOI analysis. The next section presents the error correlations of the forecast which provides guess, or background, fields for the analysis. Section 3 outlines the specification of observational error characteristics for the regional scheme. Section 4 covers the univariate analysis of relative humidity. In Section 5, some design features of the scheme are discussed, and possible changes that might be made to improve it are indicated. The final section compares some MOI analysis results with those obtained operationally for the LFM.

2. Error Correlations of the Guess Field

The correlation relations between temperature and wind errors, between temperature and height errors, and between height and wind errors are derived in Appendices A, B (Part 1), and B (Part 2) respectively of Office Note 200. These correlation functions are based on the assumption that the thermal wind relation holds between temperature and wind residuals, the hydrostatic relation applies between temperature and height residuals, and that the geostrophic equation relates height and wind residuals.

As in Office Note 200, assume that the three-dimensional correlations are products of lateral correlations (at constant pressure) and vertical correlations, i.e.,

$$\rho_{ij}^{kl} = \mu_{ij}^{kl} v_{ij}^{kl} . \quad (2.1)$$

Additionally, the following forms (eq. 3.18 and 3.19 of Office Note 200) are assumed for two of the correlations:

$$\mu_{ij}^{tt} = \mu_{ij}^{zz} = \exp[-k_h (\Delta s_{ij})^2], \quad (2.2)$$

where Δs_{ij} is the separation distance between points i and j, and

$$v_{ij}^{uu} = v_{ij}^{zz} = \frac{1}{1 + k_p \ln^2(p_i/p_j)}, \quad (2.3)$$

where p_i and p_j are the pressures of points i and j, respectively.

The analysis is done on a Cartesian grid on a polar-stereographic projection. Hence

$$\Delta s_{ij} \equiv \frac{2}{m_i + m_j} [(x_i - x_j)^2 + (y_i - y_j)^2]^{1/2}, \quad (2.4)$$

$$m_i = \frac{1}{1 + \sin \phi_i}, \quad (2.5)$$

x and y are right-orthogonal coordinates on the Cartesian grid, and u and v are horizontal wind components corresponding to directions x and y.

With (2.2) and (2.3) assumed, the remaining correlations between the residual quantities t, u, v, and z are:

$$\mu_{ij}^{ut} = -\mu_{ij}^{uz} = -\frac{(2k_h)^{1/2}}{\overline{m}} (y_i - y_j) \mu_{ij}^{tt}, \quad (2.6)$$

$$\mu_{ij}^{tu} = -\mu_{ij}^{zu} = \frac{(2k_h)^{1/2}}{\overline{m}} (y_i - y_j) \mu_{ij}^{tt}, \quad (2.7)$$

$$\mu_{ij}^{vt} = -\mu_{ij}^{vz} = \frac{(2k_h)^{1/2}}{\overline{m}} (x_i - x_j) \mu_{ij}^{tt}, \quad (2.8)$$

$$\mu_{ij}^{tv} = -\mu_{ij}^{zv} = -\frac{(2k_h)^{1/2}}{\overline{m}} (x_i - x_j) \mu_{ij}^{tt} \quad (2.9)$$

$$\mu_{ij}^{uu} = [1 - \frac{2k_h}{\bar{m}^2}(y_i - y_j)^2] \mu_{ij}^{tt}, \quad (2.10)$$

$$\mu_{ij}^{vv} = [1 - \frac{2k_h}{\bar{m}^2}(x_i - x_j)^2] \mu_{ij}^{tt}, \quad (2.11)$$

$$\mu_{ij}^{uv} = \mu_{ij}^{vu} = \frac{2k_h}{\bar{m}^2}(x_i - x_j)(y_i - y_j) \mu_{ij}^{tt}, \quad (2.12)$$

$$\mu_{ij}^{zt} = \mu_{ij}^{tz} = - \mu_{ij}^{tt}, \quad (2.13)$$

$$v_{ij}^{vv} = v_{ij}^{uv} = v_{ij}^{vu} = v_{ij}^{uu}, \quad (2.14)$$

$$v_{ij}^{ut} = v_{ij}^{vt} = v_{ij}^{zt} = (2k_p)^{\frac{1}{2}} \ln(p_i/p_j) (v_{ij}^{uu})^2, \quad (2.15)$$

$$v_{ij}^{tu} = v_{ij}^{tv} = v_{ij}^{tz} = - (2k_p)^{\frac{1}{2}} \ln(p_i/p_j) (v_{ij}^{uu})^2, \quad (2.16)$$

$$v_{ij}^{tt} = [1 - 4k_p \ln^2(p_i/p_j) v_{ij}^{uu}] (v_{ij}^{uu})^2. \quad (2.17)$$

In (2.6) through (2.12), $\bar{m} \equiv (m_i + m_j)/2$.

The above correlations are the ones currently used in the regional MOI scheme. The current value of k_h ranges between 1.0×10^{-6} and $2.0 \times 10^{-6} \text{ km}^{-2}$, depending on the observational density as described in Section 5. The current value of k_p is 2.5 (dimensionless). These values are loosely based on forecast error correlation statistics.

A result of the above partitioning of the correlations is that the mean square residuals t , u , v at point i are (in theory) related to the mean square residual z_i as follows:

$$\overline{t_i^2} = \frac{2k_p g^2}{R^2} \overline{z_i^2}, \quad (2.18)$$

$$\overline{u_i^2} = \overline{v_i^2} = \frac{2k_h g^2 G_i^2}{f_i^2} \overline{z_i^2}, \quad (2.19)$$

where g is the acceleration due to gravity, R is the meteorological gas constant, f is the Coriolis parameter, and G , the "coefficient of geostrophy," is defined as

$$G_i \equiv 1 - \exp(-.05 |\phi_i|), \quad (2.20)$$

where ϕ_i is the latitude in degrees of point i .

Substitution of the above mean square residuals in (1.8) gives the following relations between the scaled weights a and the unscaled weights a' :

(1) For height analysis:

$$a_{iz} = a'_{iz} \quad (2.21)$$

$$a_{it} = \frac{R}{g\sqrt{2k_p}} a'_{it} \quad (2.22)$$

$$a_{iu,v} = \frac{f_i}{g G_i \sqrt{2k_h}} a'_{iu,v} \quad (2.23)$$

(2) For temperature analysis:

$$a_{it} = a'_{it} \quad (2.24)$$

$$a_{iz} = \frac{g\sqrt{2k_p}}{R} a'_{iz} \quad (2.25)$$

$$a_{iu,v} = \frac{f_i}{R G_i} \sqrt{\frac{k_p}{k_h}} a'_{iu,v} \quad (2.26)$$

(3) For wind component analysis:

$$a_{iu,v} = \frac{f_i G_i}{f_i G_i g} a'_{iu,v} \quad (2.27)$$

$$a_{iz} = \frac{g G_i \sqrt{2k_h}}{f_i g} a'_{iz} \quad (2.28)$$

$$a_{it} = \frac{R G_i g}{f_i g} \sqrt{\frac{k_p}{k_h}} a'_{it} \quad (2.29)$$

In obtaining (2.21) through (2.29) the mean square residual has been assumed "locally constant."

Gandin (personal communication) has suggested that a horizontal correlation function for temperatures and heights due to Yudin,

$$\mu_{ij}^{tt} = \mu_{ij}^{zz} = \exp(-k_h \Delta s_{ij})(1 + k_h \Delta s_{ij}), \quad (2.30)$$

may be a better choice than the Gaussian one, eq. (2.2), especially in those areas where the analysis is actually extratropolating, rather than interpolating, meteorological fields.

Eqs. (2.18) and (2.19) indicate that theoretical relationships exist between the mean-square residual values which are at most a function of latitude only. How "realistic" are these relations? To find out, the rms differences between observed and 12-hr hemispheric prediction values of height, temperature, and wind speed were obtained (from J. Stackpole) for May 1979. The ratios of these differences for four pressure levels were computed and are given below:

<u>RMS Ratio of</u>	<u>850</u>	<u>500</u>	<u>250</u>	<u>100 mb</u>
t to z (deg m ⁻¹)	.118	.062	.058	.042
s to z (sec ⁻¹)	.186	.179	.154	.058

These statistics were not available as a function of latitude. Clearly, they show some variation with pressure.

From (2.18), with $k_p = 2.5$,

$$(\overline{t_i^2/z_i^2})^{1/2} = .076 \text{ deg m}^{-1} \quad (2.31)$$

From (2.19), with $k_h = 1.4 \times 10^{-12} \text{ m}^{-2}$ and with latitude $\phi_i = 45$ degrees,

$$(\overline{s_i^2/z_i^2})^{1/2} = .142 \text{ sec}^{-1} \quad (2.32)$$

How serious is this discrepancy between "reality" and "theory"? Note that these ratios determine the scaling factors of (2.21) through (2.29). Thus, when cross-correlated observations are used in the analysis of a variable, they are "overscaled" or "underscaled" to some extent for some pressure levels. A possible "fix" for this situation is to let k_h and/or k_p be a function of the pressure level of the analysis. This experiment has not yet been tried. Also, k_h and/or k_p can be assigned different values when analyzing different variables. In the global MOI analysis, k_p is assigned a different value when t is analyzed from that used when u or v are analyzed. The regional OI analysis code is designed to accept three different values of k_p for the analysis of temperature, height, and wind components, but thus far only one value of k_p has been used for all three analyses.

3. Specification of Observational Error Characteristics

Eq. (1.3) contains error correlation terms in which the observational error correlation, η , and the rms observational error, ϵ , appear. When $i=j$ and $k=\ell$ in (1.3), the observational error term reduces to ϵ_{ik}^2 , the normalized mean square error of observational type k which is located at point i . Normalization is with respect to the mean square error of the guess value. Table 1 lists the values of ϵ_{ik}^2 currently used in the regional MOI analysis.

When $k \neq \ell$ in (1.3), observational errors are uncorrelated, hence $\eta = 0$. When $k = \ell$ but $i \neq j$, the observed data may have correlated errors if the data were obtained sequentially by the same instrument. Vertically correlated errors of rawinsonde serial ascents have been demonstrated by Hollett (1975). Laterally as well as vertically correlated errors of temperatures derived from satellite radiances have been shown by several investigators (see Bergman, 1978). These error correlations are modeled in the analysis code, and the vertical correlations are shown in Fig. 1. The lateral correlation of satellite temperature errors is modeled by

$$(\eta_{ij})_h = \exp[-\kappa \Delta s_{ij}^2] , \quad (3.1)$$

where $\kappa = 11.3 \times 10^{-6} \text{ km}^{-2}$ and Δs_{ij} is the observational separation distance. This functional form is modeled on NIMBUS satellite data; preliminary computations show approximately the same spatial error correlation for TIROS-N satellite temperatures (Fig. 2). The above error correlation models are the same (at this writing) as those used in the global MOI scheme.

There is no information currently available on the error correlations of aircraft or satellite wind data. In the absence of information, these are assumed to be zero by the analysis code.

4. Relative Humidity Analysis

The relative humidity analysis is univariate, and it is done for the midpoints of the three lowest sigma-layers of the LFM model. In this respect, it parallels the currently operational relative humidity analysis for the LFM model (Chu, 1977). The data base is the same as for the currently operational analysis and includes sigma-layer RH values inferred from surface cloud and weather reports.

The RH analyses for the sigma layers of a grid point column are not done until after the temperature, height, and wind component analyses have been done for the mandatory levels of the same column. The results of the mass and momentum analyses are used to determine the form of the correlation function for each sigma-layer RH analysis. Thus, although the RH analysis is mathematically univariate, indirect use is made of mass and momentum information.

The analysis is performed at those grid points located at the intersections of odd-numbered grid lines (Fig. 2). Thus, the analysis is on twice the LFM-I grid. After the analysis is completed on this grid, it is bi-quadratically interpolated to the remaining LFM-I grid points.

Although the LFM model now runs on the LFM-II grid, the "initializer" and graphics codes currently expect the analysis on the LFM-I grid. The analysis code can be modified for analysis on any Cartesian grid on a polar-stereographic projection. The analysis code and attendant data preparation routines require about 10 min CPU time for analysis on the 27 x 29 grid at 10 mandatory pressure levels (1000 through 100 mb) and, for RH, three sigma levels. Some code changes to improve time efficiency are currently being tried in the global scheme and, if successful, will also be incorporated in the regional scheme.

b. The Guess, or Background, Field

Currently, the 12-hr hemispheric PE numerical prediction provides the guess fields for the regional MOI analysis. However, the forecast wind fields are blended with a wind field computed from the forecast heights using the form of the gradient wind equation given in McDonnell (1973). The blending is 50% forecast winds and 50% gradient winds. The motivation behind this is to provide guess wind fields in somewhat better balance with the height fields than are provided by the forecast winds. Limited testing indicated some overall improvement in the analysis when the blended guess winds were used compared to 100% forecast winds.

A possible extension of this idea is to blend the forecast heights with heights computed from the forecast winds via the inverse balance equation. This experiment has not been tried.

c. Use of Surface Data in the Upper-Air Analysis

Reported sea-level pressures and oceanic ship winds are used in the upper-air analysis. The sea-level pressures are converted to equivalent 1000-mb geopotential heights, and the ship winds are converted to approximate equivalent geostrophic winds. The procedure and equations for doing this are described in Section 6 of Office Note 200. The converted ship winds are treated as 1000-mb wind data by the analysis code. Land wind reports are not used. A separate univariate surface temperature analysis is performed (see below), but surface temperature data is not used in the upper-air analysis. For rawinsonde reports when both sea-level pressures and 1000-mb heights are available, the conversion of pressure to equivalent height is suppressed.

e. Computation of Analysis Weights and Field Corrections

The observational weights (the a_{jl} of eq. 1.3) are computed precisely as in the global scheme (see Office Note 200, Sec. 7b). The scale factors of Section 2 are then used to rescale the weights for cross-correlated data. The analyzed correction for each of the four variables at each level of each grid point is then given by (1.1). The corrections are added to the guess values to get the analyzed gridded fields.

In some experimental analyses, the guess values were adjusted individually at each grid point and level for "bias" by first averaging the residuals (if 5 or more) of all data autocorrelated with the variable being analyzed. This average residual was then assumed to be the bias in the guess variable, and it was algebraically added to the guess and subtracted from each of the autocorrelated residuals. In practice, this procedure produced some undesirable irregularities in the analyzed fields and was discontinued, although the coding to do it is still present in the analysis program.

Needless to say, if no observations are found by the search routine, the correction will be zero and the guess value remains unaltered.

f. Surface Temperature Analysis

A univariate surface temperature analysis is performed. The lateral correlation function for upper-air temperature analysis is used unchanged in the surface analysis. (There is no requirement that this be the case.) By accident, the bias correction described in the preceding subsection was included. It appears to work favorably here, probably because the forecast surface temperatures often differ so markedly from the reported values.

g. Tropopause Analysis

The LFM initializer requires tropopause temperature, potential temperature, and pressure analyzed fields. As of this writing, no optimum interpolation analysis code has been written for the tropopause. Immediate plans are to incorporate the presently operational tropopause analysis as a "black box" in the regional MOI analysis package. In tests of the MOI scheme on archived data sets, the operationally produced tropopause analysis file was added to the other files produced by the OI scheme.

h. Data Quality Control

The gross and comparative error checking is done in the same way as described in Office Note 200, Section 7c. The gross error check is done in terms of the forecast error standard deviations given in Table 2. "A" quality rawinsonde reports within 5 of the guess value are accepted, otherwise they are rejected. All other observations have an acceptance threshold of 3. Relative humidity reports are rejected only if their values lie outside the range of 0 to 100%.

i. Smoothing-Desmoothing the Analysis

When the analysis is performed on twice-the-LFM-I grid and bi-quadratically interpolated to the rest of the grid points, it contains some small-scale oscillations which are visually unattractive and which are not always indicated by the observed data. An example is shown in Fig. 3. Application of the Shuman (1978) nine-point smooth-desmooth filter with coefficients $v = \frac{1}{2}, -\frac{1}{2}$ produces the analysis of Fig. 4. This analysis is comparable in smoothness to the operationally produced analysis for the same time, Fig. 5. It turns out, however, that the LFM forecasts out to 48 hours from the filtered and unfiltered analyses were almost identical. The filtering would appear to be of cosmetic value only.

The analysis code contains a switch (in POLA) that allows the smooth-desmooth filter to be turned on or off.

6. Some Regional MOI Analysis Results

Several test analyses have been done. These have been used to initialize the LFM model, from which 48-hr numerical predictions have been made. Thus far, these have only been evaluated subjectively. Objective, as well as subjective, evaluation of analyses and forecasts, including comparison with the currently operational LFM analysis and with R. Jones's analysis in isentropic coordinates, is planned by members of Atmospheric Analysis Branch. The LFM model version of D. Deaven with improved convection will be used for forecast comparisons.

Thus far, the MOI scheme appears to delineate the atmospheric circulation with about the same fidelity as the successive corrections scheme. The smoothed versions do not contain the objectionable high-frequency oscillations in the data-sparse areas that the unsmoothed versions frequently show. On the other hand, the smoothed version does not fit the observations so closely as the unsmoothed version. Compare, for example, Tables 3 and 4. The comparable statistics for the operational LFM analysis are not available. Closeness of fit to the observations is not the primary goal of the analysis, nor is it a particularly good measure of analysis quality. See the discussion in Office Note 200 on this point.

LFM numerical predictions based on the MOI analysis generally show circulation patterns which are very similar to those resulting from the currently operational analysis. The primary difference in forecasts appears to be in the moisture and precipitation forecasts. The MOI scheme has shown some ability to increase the forecast precipitation in areas where forecasts based on the operational analysis are deficient. This is well illustrated by the predictions of the snowstorm of 18-19 February 1979, discussed below.

At 1200Z 18 February 1979, the surface chart showed a large anticyclone centered near Lake Ontario and an inverted trough in the lower Mississippi Valley, with a frontal wave forming along the Gulf Coast (Fig. 5). At 500 mb,

a double trough was located over the Northern Plains and Mississippi Valley (Fig. 4). Twenty-four hours later, the anticyclone had moved off the east coast, a rapidly deepening cyclone was just east of the Virginia Capes (Fig. 10), and the 500-mb trough had moved eastward with its vorticity maximum over West Virginia.

The operational LFM surface analysis (after initialization) is shown in Fig. 6, and the comparable OI analysis (smoothed version) is shown in Fig. 7. These are very similar. The mean RH analyses for the three lowest model layers, superimposed on the 700-mb height analyses, are shown in Figs. 8 and 9. These also are similar in the area of interest. (Note that the RH analyses produced by the MOI scheme are not smoothed by the nine-point filter, even though all the other fields are smoothed in this version.) Of course, these charts do not show differences in the vertical profile of moisture that may exist between the two analyses.

The 24-hr surface/thickness forecasts produced by the LFM code which was operational on 18 Feb 1979 are shown in Figs. 11 and 12. The forecasts are similar, but the MOI version has the developing low closer to the coastline and has a stronger gradient of pressure north of the center. Note that the MOI version correctly forecast higher pressures in southern New England. Both versions have the low center too far south and not intense enough.

The 24-hr 700-mb/mean RH forecasts are shown in Figs. 13 and 14. Both versions show excellent forecasts of the upper trough. At this time, the mean RH forecasts are again very similar. The precipitation which occurred during the preceding 12 hrs is shown in Fig. 15, and the corresponding operational and MOI forecasts are shown in Figs. 16 and 17. The MOI version shows heavier precipitation over the Chesapeake Bay area than the operational version. It is still short of the amounts which actually fell in the area. Too much precipitation is predicted in the Carolinas, probably associated with the too-far-south location of the low and the usual slowness of the LFM model to end precipitation behind an east coast system.

Figure 18 shows the observed 12-hr precipitation in the West. Note that the MOI version (Fig. 17) has predicted light precipitation in northern California and Nevada, in contrast to the operational version (Fig. 16).

The observed precipitation for the eastern storm during the first 12 hours of the forecast period is shown in Fig. 19. The corresponding operational and MOI forecasts are displayed in Figs. 20 and 21, respectively. The MOI version is superior with respect to both the location of the maximum and the maximum amount predicted. Fig. 22 shows the observed precipitation for the 24-36 hr forecast interval. The forecast from the operational analysis is shown in Fig. 23, and that from the MOI analysis in Fig. 24. Although both forecasts seriously underestimate the precipitation in the New Jersey-Delaware area, the MOI version does indicate greater amounts here than the operational version.

When the unsmoothed version of the MOI analysis was used to initialize the LFM forecast from 12Z 18 Feb 1979, results were very similar to those obtained from the smoothed version. The 12-hr precipitation forecast for the crucial 12-24 hr period is shown in Fig. 25. It predicts slightly more precipitation over eastern Virginia and slightly stronger vertical motions offshore than does the smoothed version (Fig. 17).

Both the unsmoothed and smoothed MOI versions have been used to initialize Deaven's improved convection LFM model. In both cases, forecast precipitation amounts were even greater than shown in Fig. 25, with little difference between the unsmoothed version (Fig. 26) and the smoothed version (Fig. 27).

7. Conclusions

The design of a regional multivariate optimum interpolation analysis scheme has been described above, and analysis/forecast results for one case have been presented and discussed. This case indicates that the MOI scheme is capable of producing analyses and forecasts of comparable skill to those produced by the currently operational LFM analysis. An improvement in skill of the precipitation forecasts is shown.

Although multivariate optimum-interpolation is based on a sound theoretical framework, it is obvious that many compromises and empirical adjustments have been made in this MOI scheme in order to make it feasible for operational use. Similar compromises have of necessity been made in other OI schemes intended for operational use. Some of the deficiencies are due to ignorance; for example, the error characteristics of some kinds of observations are not well known. Other weak points are a result of the need for time-efficiency; we know how the data screening and selection procedure could be improved, but more computation time would be required. Finally, some limitations are imposed by the theory; the correlations must be modeled by analytic functions in order for the coefficients of eq. (1.3) to define a positive-definite matrix (Gandin, 1963). All these compromises have led some to call such schemes "pseudo-optimum interpolation" (and occasionally even worse names).

Is the departure of the actual analysis scheme from the theoretical ideal a serious limitation? Note that so-called "optimum interpolation" schemes are only optimized in a statistical sense, even in theory. Correlations are based on the historical performance of forecasts used for guess values and on the statistical error characteristics of large ensembles of observations. The guess forecast and observational errors for individual analyses may differ markedly from the norm. In practice, one must regard the MOI scheme as just another device for spreading the observed information to the grid points and making some compensatory mass-momentum balance adjustments. As much empiricism as there is in the scheme, it can still be related to a theoretical framework capable of suggesting future improvements as well as providing a fairly exact measure of what compromises have been made in the existing version. Many other analysis schemes are strictly products of trial and

error. Ultimately, the uses to which an analysis scheme is put, be they visual display or initializing forecast models, must determine whether or not an analysis system is a "good" one, compared to its competitors.

REFERENCES

- Bergman, K. H., 1978: Role of observational errors in optimum interpolation analysis. Bulletin, AMS, 59, 1603-1611.
- _____, 1979: Multivariate analysis of temperature and wind fields using optimum interpolation. Office Note 200, National Meteorological Center. (Submitted to Monthly Weather Review)
- Chu, R., 1977: Humidity analyses for operational prediction models at the National Meteorological Center, Part I: LFM. Office Note 140, National Meteorological Center.
- Gandin, L. S., 1963: Objective Analysis of Meteorological Fields. In Russian. Gidrometeorologicheskoe Izdatelstvo (GIMIZ), Leningrad. (English translation, Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.)
- _____, and R. L. Kagan, 1974: Construction of a system for objective analysis of heterogeneous data based on the method of optimal interpolation and optimum agreement. Meteoro. Gidrol., No. 5, 1-11. (English translation, Joint Publications Research Service, Arlington, Va.)
- Hollett, S. R. (1975): Three-dimensional spatial correlations of PE forecast errors. M.S. Thesis, Dept. of Meteorology, McGill University, Montreal.
- McDonnell, J. E., 1973: Notes on operational objective analysis procedures. Automation Division, National Meteorological Center.
- McPherson, R. D., et al., 1977: Global data assimilation by local optimum interpolation. Office Note 141, National Meteorological Center.
- Rutherford, I. D., 1973: Experiments on the updating of PE forecasts with real wind and geopotential data. Preprints, Third Conf. on Probability and Statistics in Atmospheric Sciences (Boulder), AMS, Boston, pp. 198-201.
- Schlatter, T. W., 1975: Some experiments with a multivariate statistical objective analysis scheme. Monthly Weather Review, 103, 246-257.
- Shuman, F. G., 1978: Smoothing and interpolation in LFM-II output. Office Note 165, National Meteorological Center.

TABLE 1.

RATIO OF RMS OBSERVATIONAL ERROR TO RMS ERROR OF GUESS FIELD

	<u>u,v</u>	<u>T</u>	<u>Z</u>	<u>RH</u>
Rawinsonde	.4	.5	.5	.25
Aircraft	.4	.5	.5	.25
Satellite	.5	1.0	1.0	1.0
Sfc. Cl. Wea.	--	--	--	.5

TABLE 2.

FORECAST ERROR STANDARD DEVIATIONS

<u>PRESSURE</u> (mb)	<u>TEMP.</u> ($^{\circ}\text{C}$)	<u>U</u> (ms^{-1})	<u>V</u> (ms^{-1})	<u>HEIGHT</u> (m)
1000	4.3	4.9	4.6	50
850	3.2	4.2	4.1	24
700	2.2	4.3	4.0	28
500	2.0	4.9	4.7	33
400	2.1	5.7	5.6	36
300	2.6	7.1	6.7	41
250	3.4	7.5	7.2	45
200	3.6	7.3	7.0	50
150	3.1	6.7	6.4	55
100	3.0	5.6	5.4	61

TABLE 3: FIT OF MOI ANALYSIS (UNSMOOTHED) TO DATA.
122, 18 FEB 1979

RMS TEMPERATURE (CENTIGRADE)									
	850	700	500	400	300	250	200	150	100
RADIOSONDE	1.26	0.62	0.64	0.49	0.64	0.79	0.74	0.52	0.53
AIRCRAFT	125.	126.	126.	115.	108.	102.	101.	100.	91.
VTPR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TWERLE	1.23	0.66	0.93	1.08	0.94	1.12	1.06	0.31	0.56
NIMBUS	4.	3.	4.	4.	4.	4.	4.	4.	4.
TOTAL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	1.07	0.63	0.65	0.52	0.65	0.80	0.76	0.51	0.53
RMS HEIGHTS (METERS)									
	850	700	500	400	300	250	200	150	100
RADIOSONDE	7.81	5.48	7.41	8.61	10.88	14.22	14.12	10.32	13.21
AIRCRAFT	127.	131.	132.	123.	110.	105.	102.	100.	93.
VTPR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TWERLE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
NIMBUS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TOTAL	7.81	5.48	7.41	8.61	10.88	14.22	14.12	10.32	13.21
RMS VECTOR WIND DEVIATION (KNOTS)									
	850	700	500	400	300	250	200	150	100
RADIOSONDE	7.76	5.27	5.97	6.59	7.61	9.03	8.39	5.66	5.77
AIRCRAFT	118.	126.	127.	114.	107.	99.	97.	92.	88.
NESS WINDS	0.0	4.74	10.72	13.31	10.48	10.77	9.46	4.96	0.0
TWERLE	0.0	1.	12.	10.	127.	255.	21.	1.	0.0
TOTAL	0.0	10.90	0.0	4.95	18.78	12.87	13.15	0.0	0.0
	0.0	2.	0.0	2.	4.	8.	7.	0.0	0.0
	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	7.76	5.40	6.52	7.33	9.52	10.38	8.90	5.65	5.77
	118.	129.	139.	126.	238.	362.	125.	93.	88.

RELATIVE HUMIDITY				
	1000	850	700	500
RADIOSONDE	0.00	20.63	24.03	19.16
	0.00	122	129	79
SFC CLOUDS	BNDRY	.66666	.33333	
	11.02	10.64	10.50	
BOGUS	1254	1340	937	
	0.0	0.0	0.0	
	0	0	0	

TABLE 4: FIT OF MOI ANALYSIS (SMOOTHED) TO DATA
122, 18 FEB 1979

RMS TEMPERATURE (CENTIGRADE)									
	850	700	500	400	300	250	200	150	100
RADIOSONDE	1.50	0.83	0.86	0.66	0.82	1.00	1.01	0.61	0.63
AIRCRAFT	125.	126.	126.	115.	108.	102.	101.	100.	91.
VTPI	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TWERLE	1.17	0.53	0.85	1.13	0.80	1.15	1.23	0.30	0.49
NIMBUS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TOTAL	1.50	0.82	0.86	0.68	0.82	1.01	1.02	0.60	0.63
RMS HEIGHTS (METERS)									
	850	700	500	400	300	250	200	150	100
RADIOSONDE	8.55	7.12	10.17	12.82	14.97	17.56	17.17	14.43	16.18
AIRCRAFT	121.	131.	132.	123.	110.	105.	102.	100.	93.
VTPI	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TWERLE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
NIMBUS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TOTAL	8.55	7.12	10.17	12.82	14.97	17.56	17.17	14.43	16.18
RMS VECTOR WIND DEVIATION (KNOTS)									
	850	700	500	400	300	250	200	150	100
RADIOSONDE	9.32	6.86	8.27	9.60	10.96	11.23	10.30	6.69	6.46
AIRCRAFT	118.	126.	127.	114.	107.	99.	97.	82.	88.
NESS WINDS	0.0	11.66	0.0	6.11	21.29	14.26	13.81	0.0	0.0
TWERLE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TOTAL	9.32	6.87	8.56	10.18	11.55	11.93	10.57	6.67	6.46
RELATIVE HUMIDITY									
	850	700	500	400	300	250	200	150	100
RADIOSONDE	0.70	20.63	24.03	19.16					
SFC CLOUDS	ENDRY	66666	33333						
BOGLS	11.02	10.64	10.50						
	1254	1340	937						
	0.0	0.0	0.0						

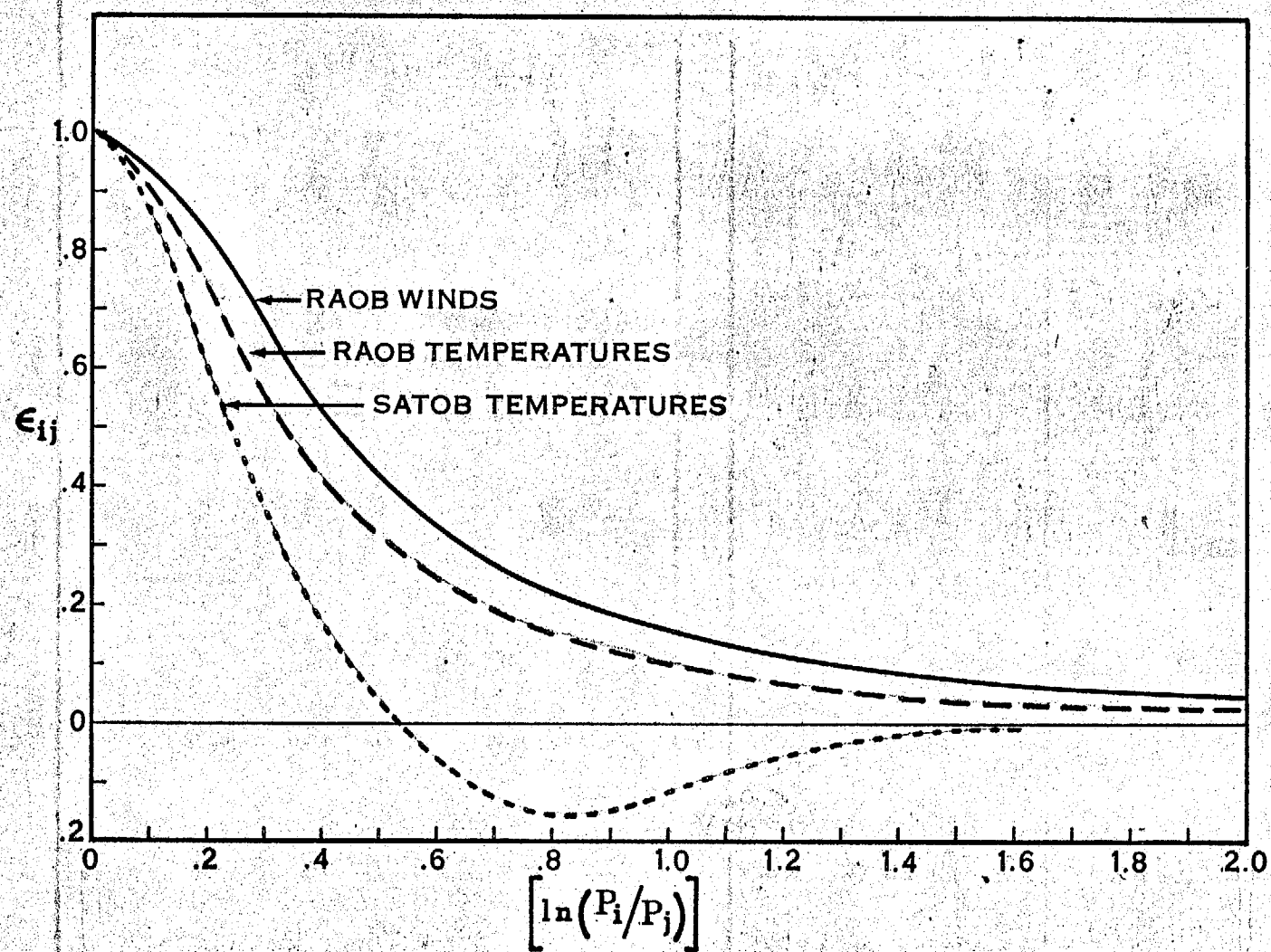


Fig. 1. Vertical Observational Error Correlations

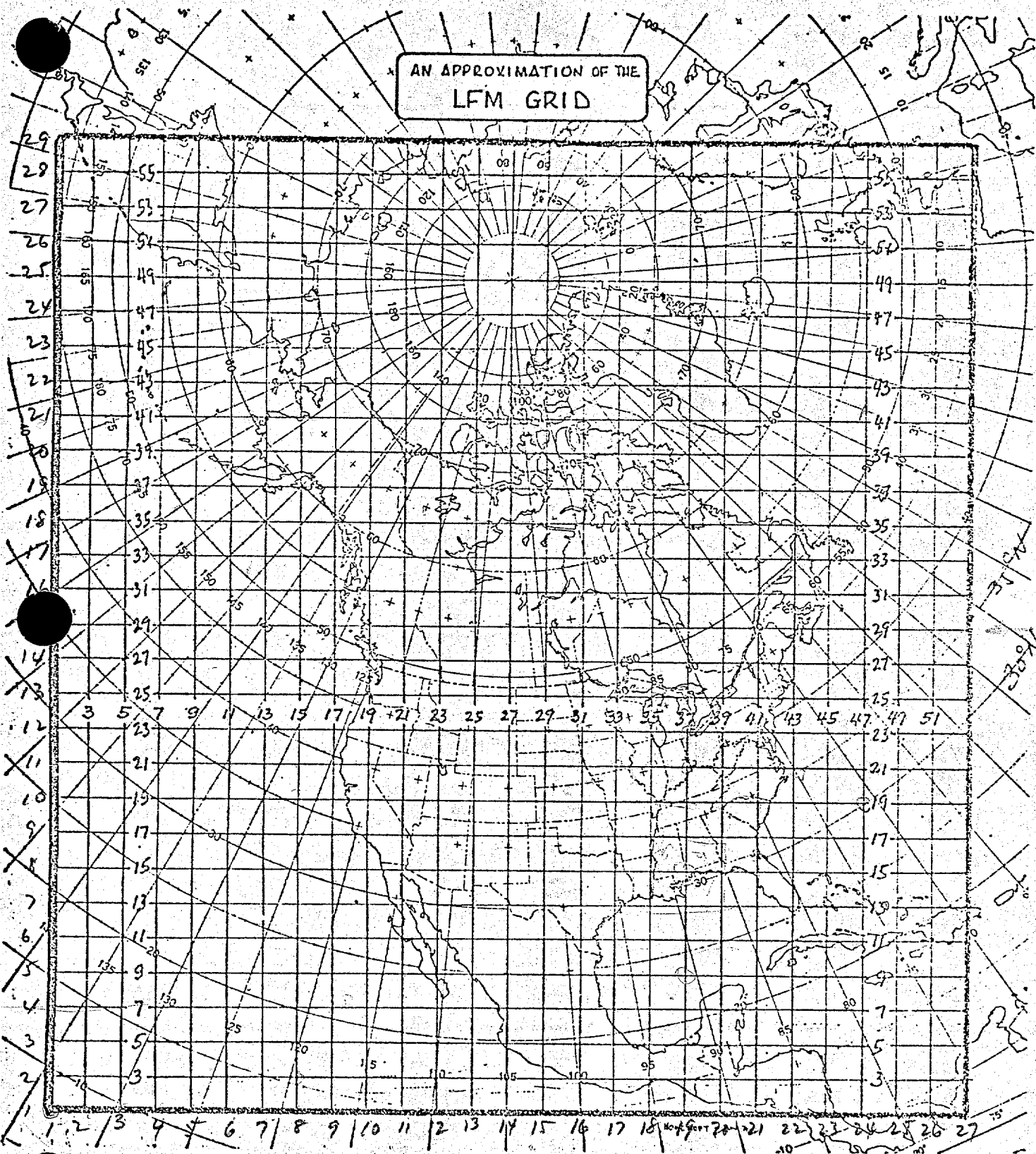
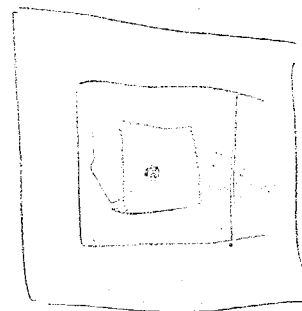


Fig. 2. The Regional MoI Analysis Grid

$$\mu_{tt} = \mu_{zz} = C - \frac{C_h}{h} S^2$$



$$\frac{1}{\psi_{uu}^{ij}} = \frac{1}{1 + C_p \ln \left(\frac{P_i}{P_{j_i}} \right)}$$



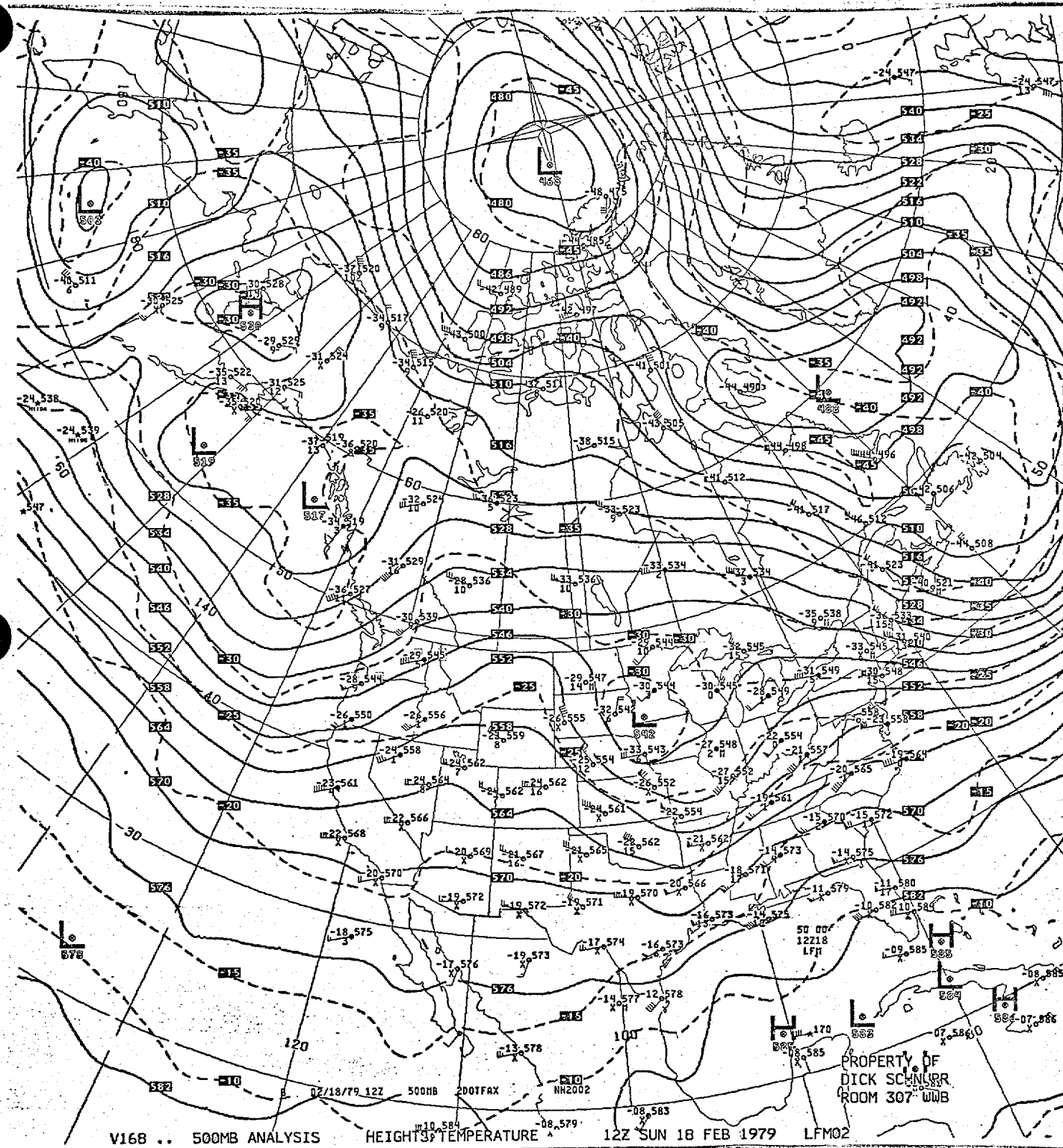
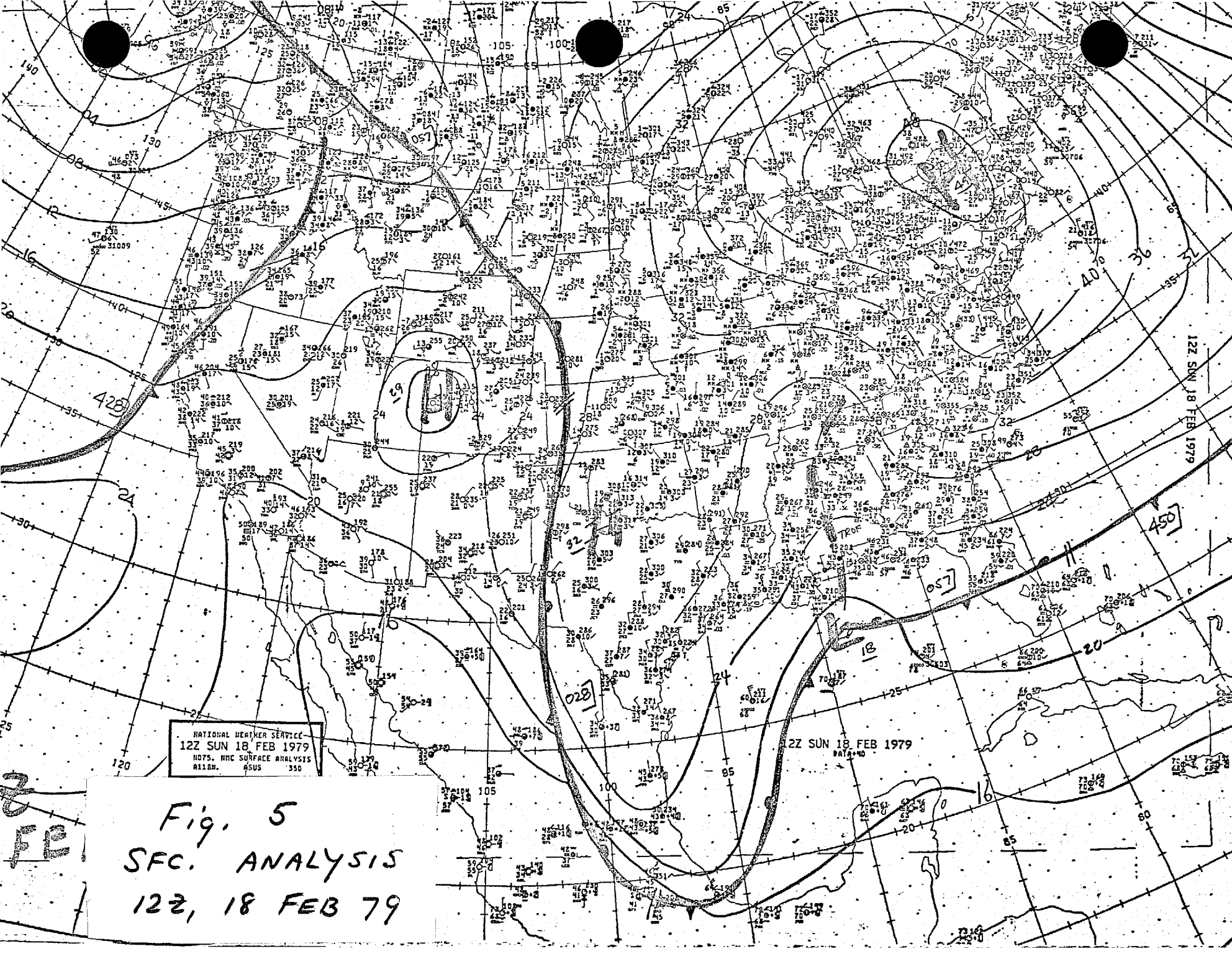


Fig. 3. UNSMOOTHED MOI. ANALYSIS



NATIONAL WEATHER SERVICE
12Z SUN 18 FEB 1979
0075, MHC SURFACE ANALYSIS
0110N, 05US, 350

Fig. 5
SFC. ANALYSIS
12Z, 18 FEB 79

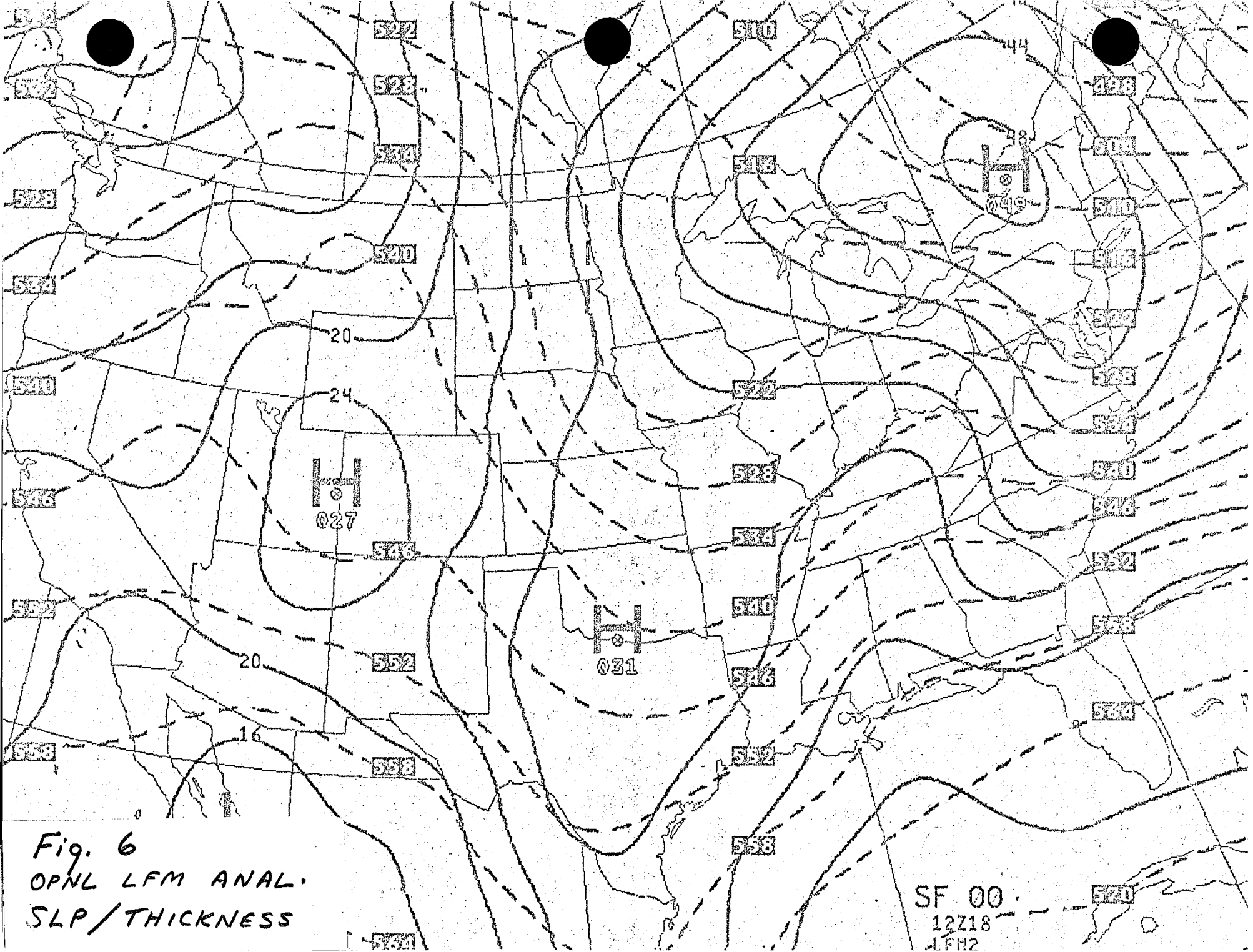
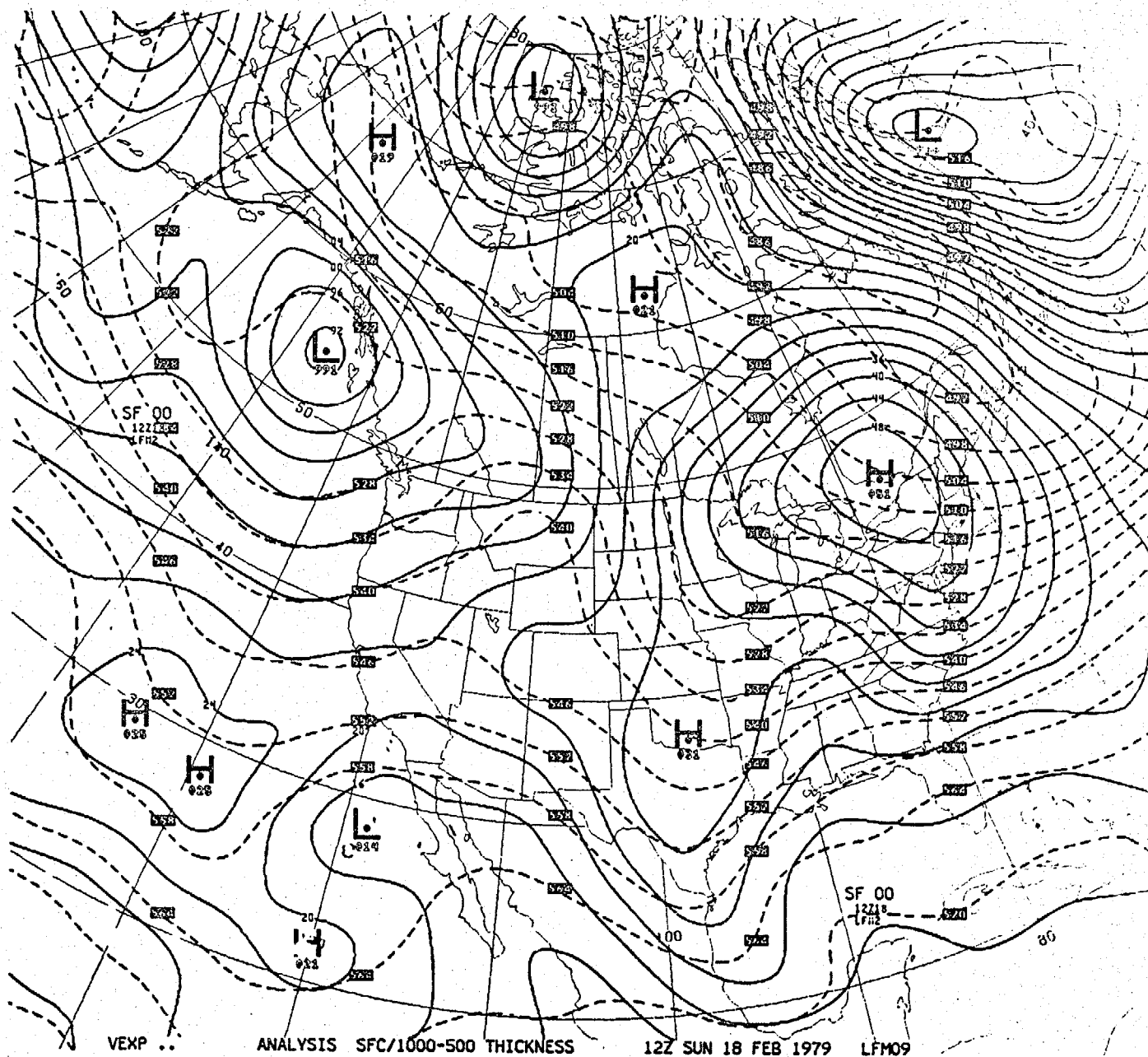


Fig. 6
OPNL LFM ANAL.
SLP/THICKNESS

SF 00
12Z18
LFM2



~~DI (F)~~ Fig. 7. MOI ANALYSIS (SMOOTHED)

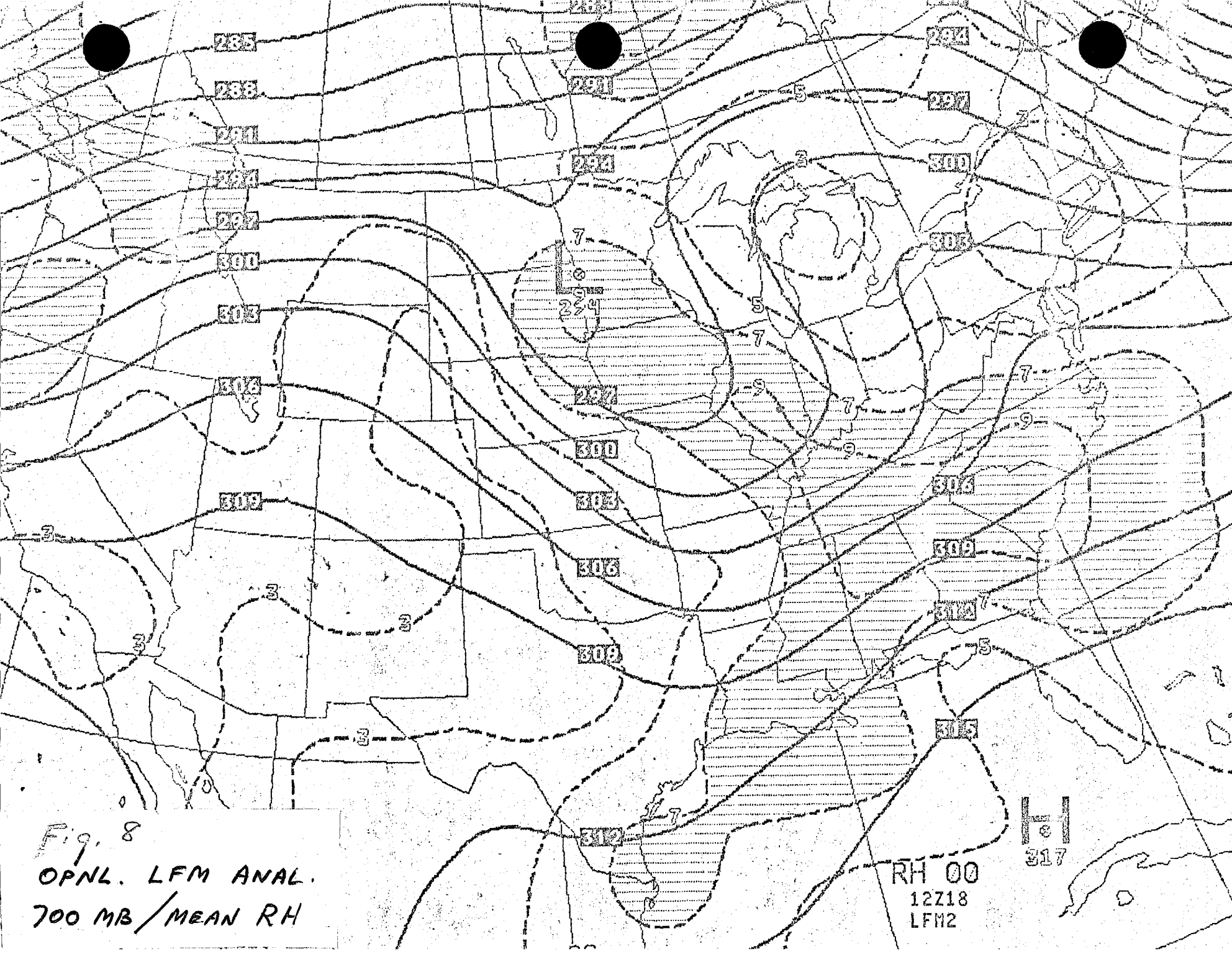


Fig. 8

OPNL. LFM ANAL.

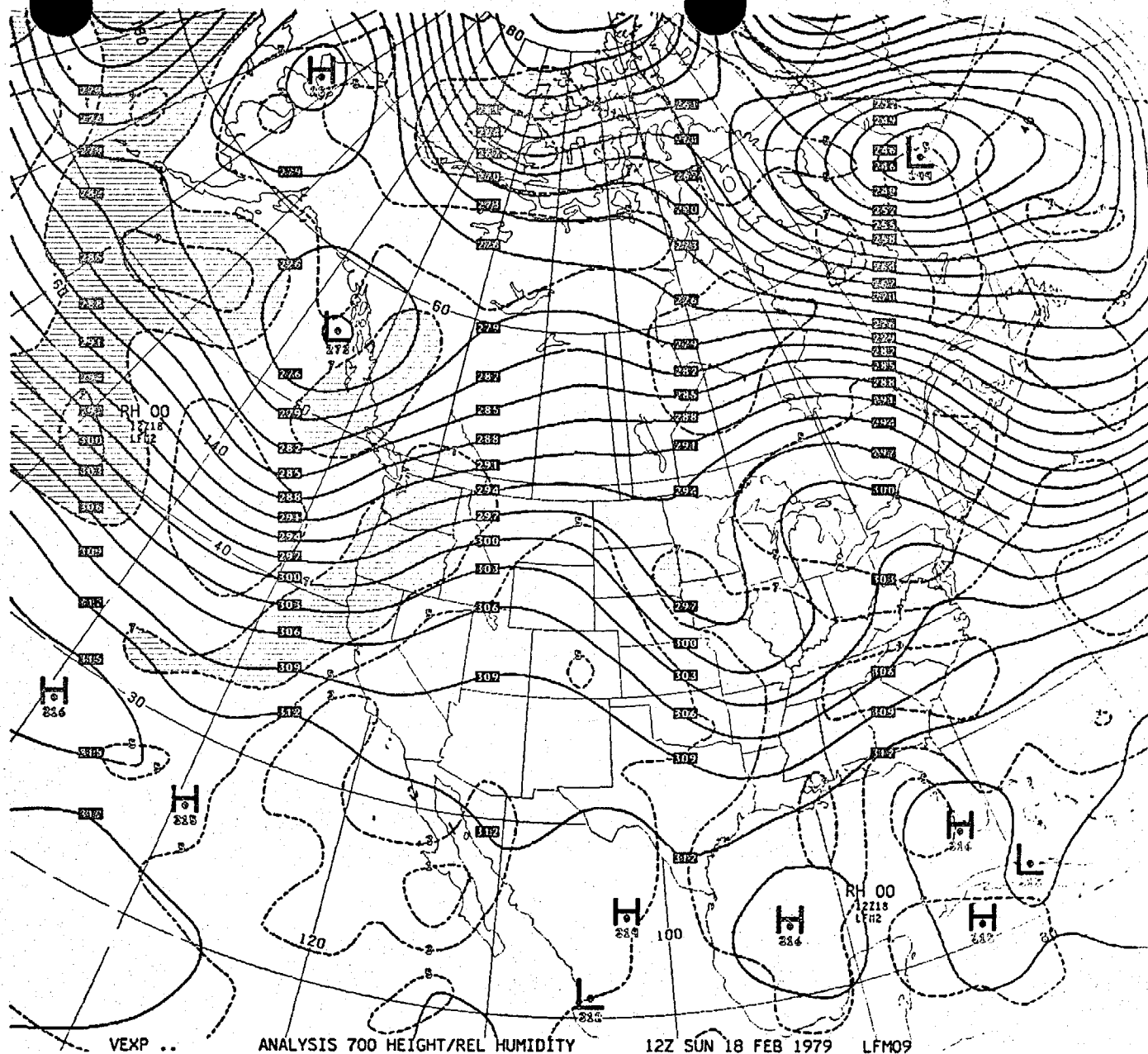
700 MB / MEAN RH

RH 00

12Z18

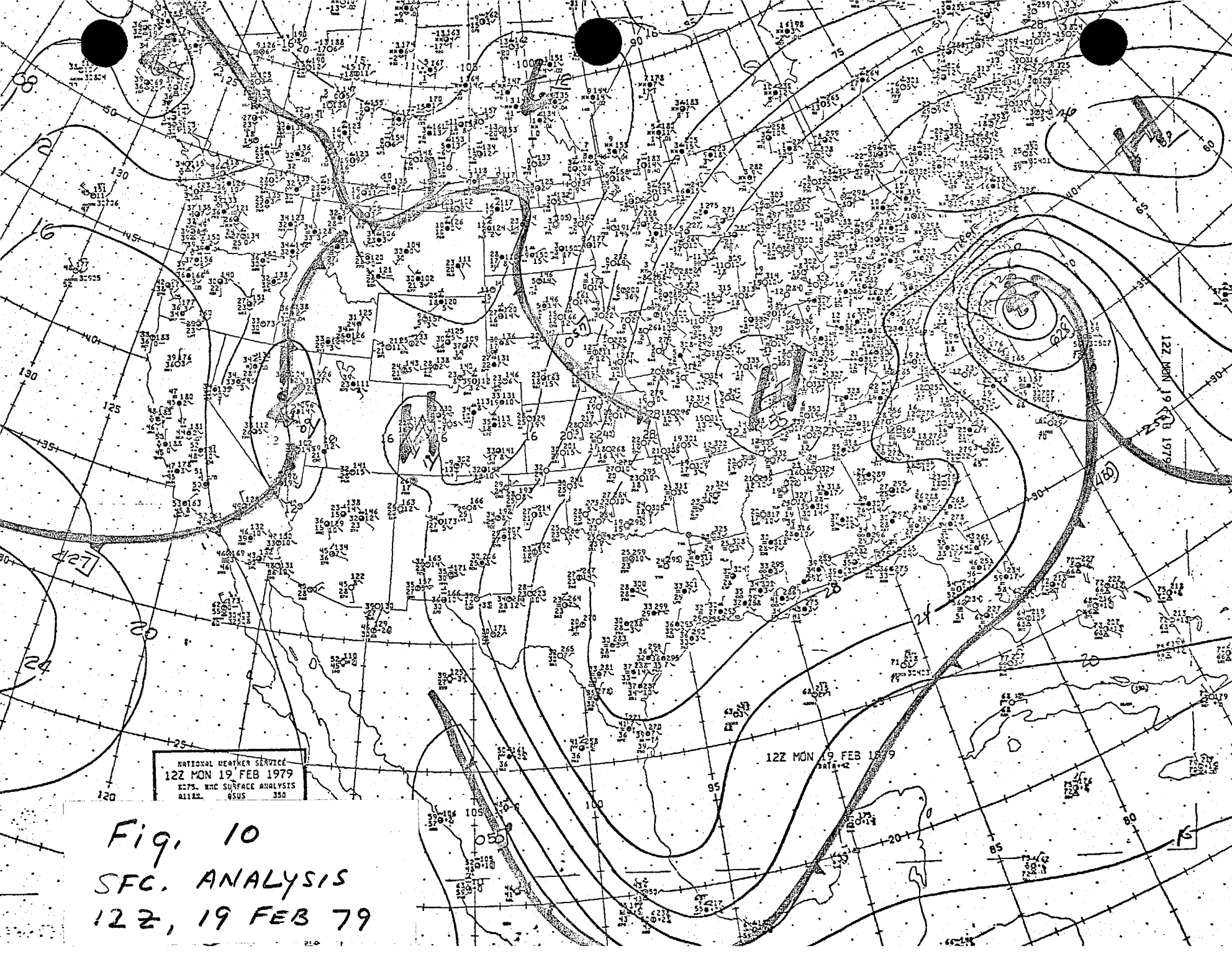
LFM2

317



01(F)

Fig. 9. MOI ANALYSIS (SMOOTHED)



NATIONAL WEATHER SERVICE
12Z MON 19 FEB 1979
KOPS. KNC SURFACE ANALYSIS
05US 350

12Z MON 19 FEB 1979
05US 350

Fig. 10
SFC. ANALYSIS
12Z, 19 FEB 79

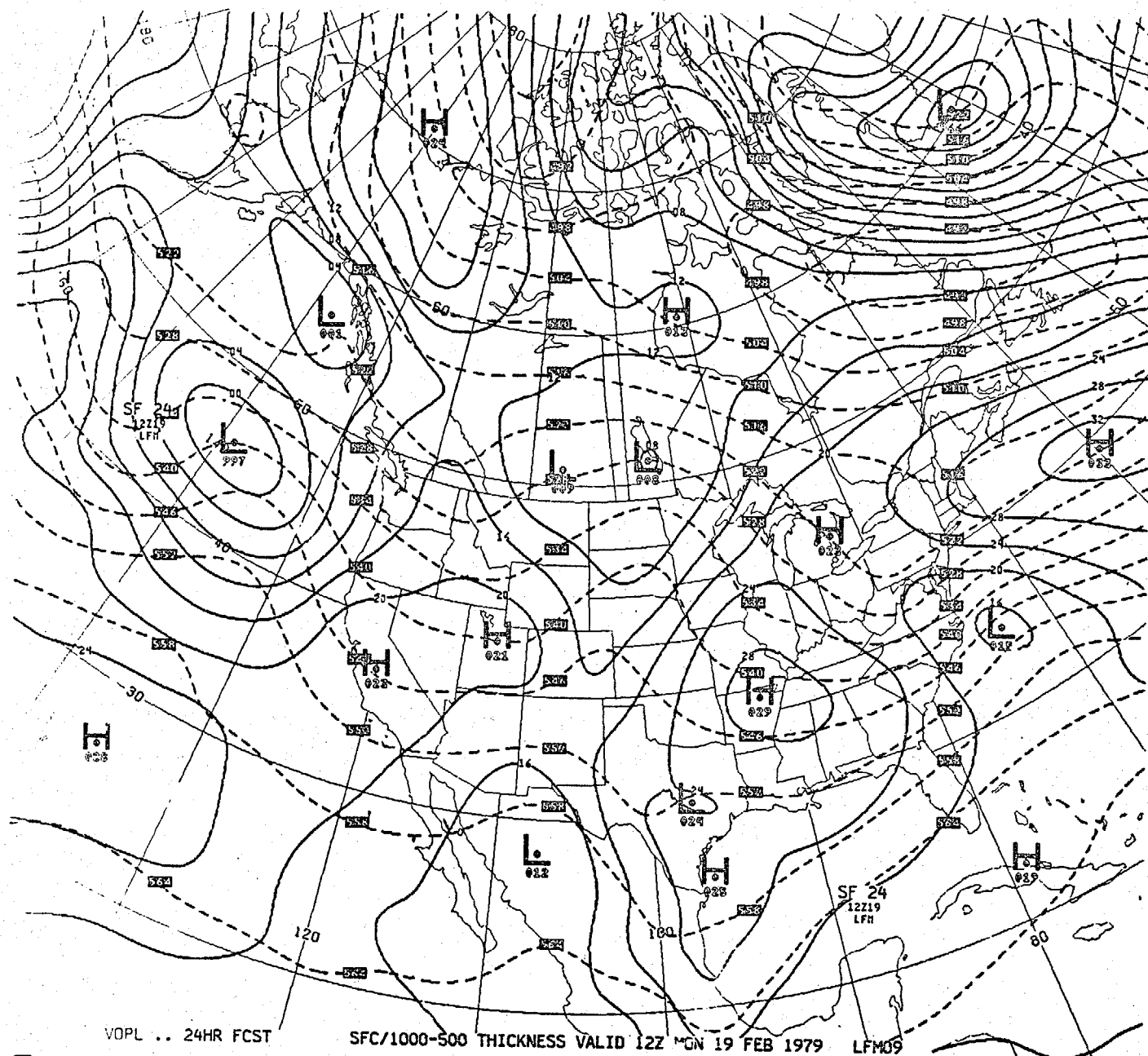


Fig. 11. LFM FCST. FROM OPNL. LFM ANAL.

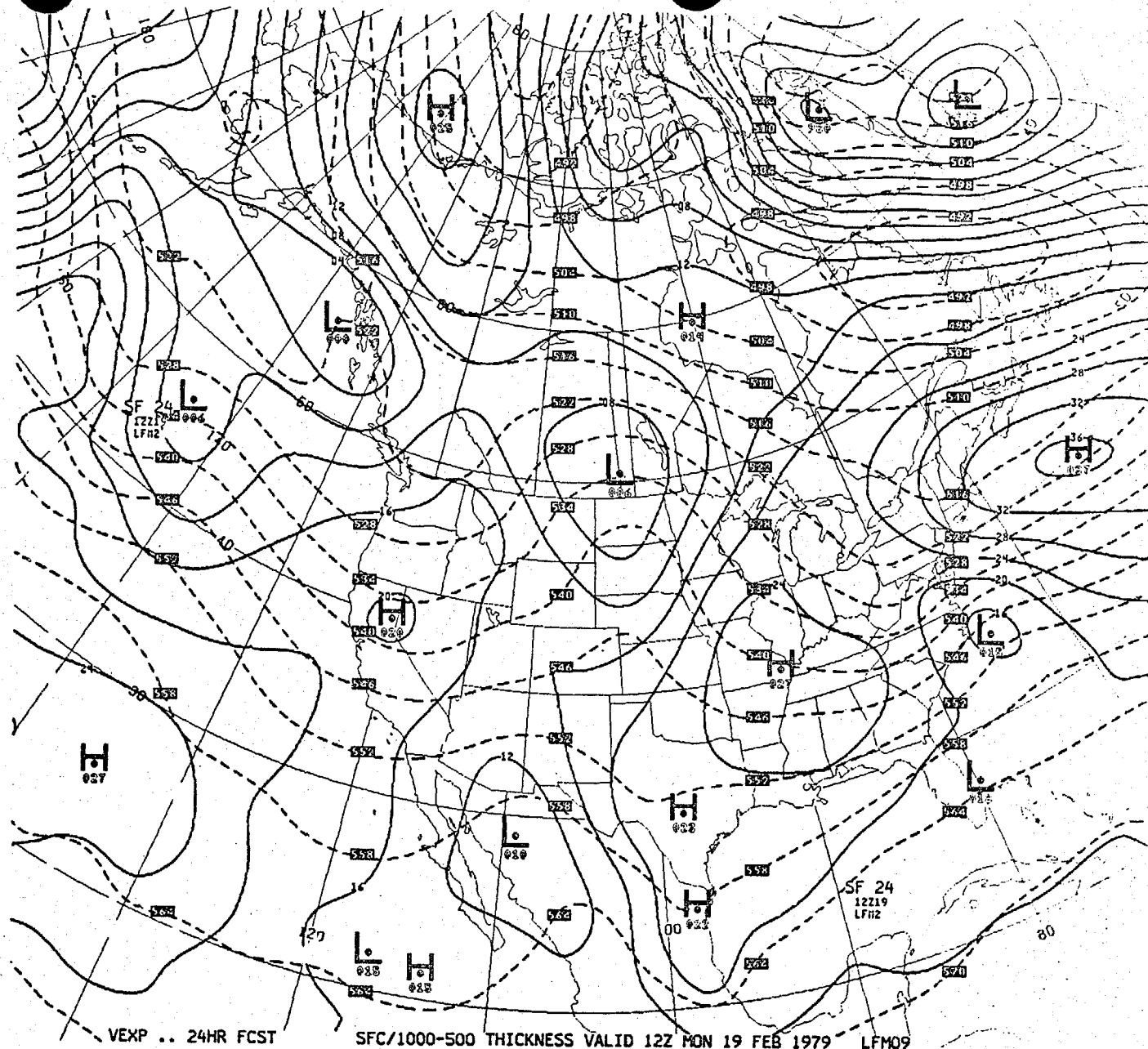
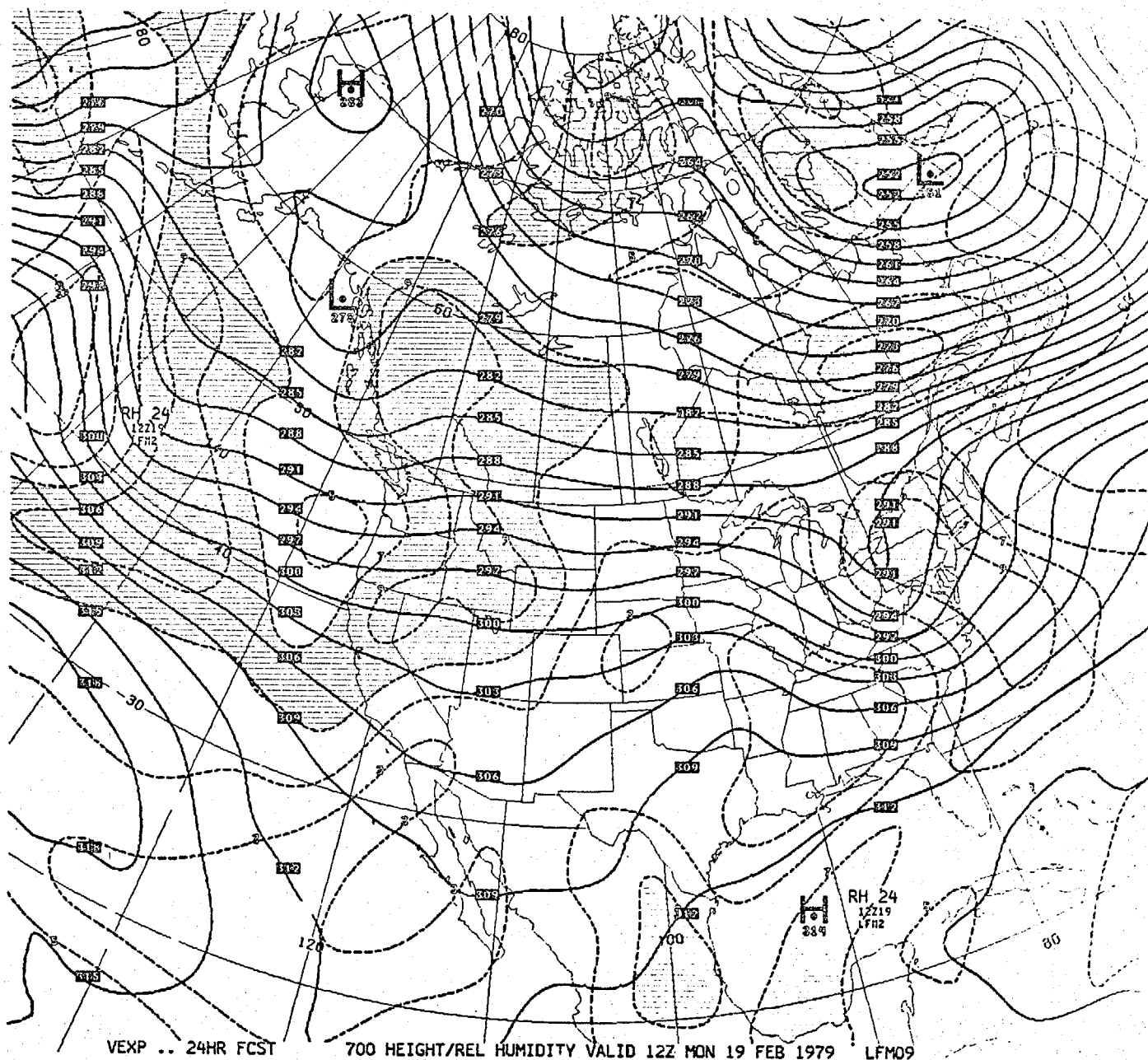


Fig. 12. ~~OI (F)~~ LFM FCST. FROM MOI ANAL. (SMOOTHED)



VEXP .. 24HR FCST

700 HEIGHT/REL HUMIDITY VALID 12Z MON 19 FEB 1979 LFM09

Fig. 14. LFM FCST FROM MOI ANAL. (SMOOTHED)

13

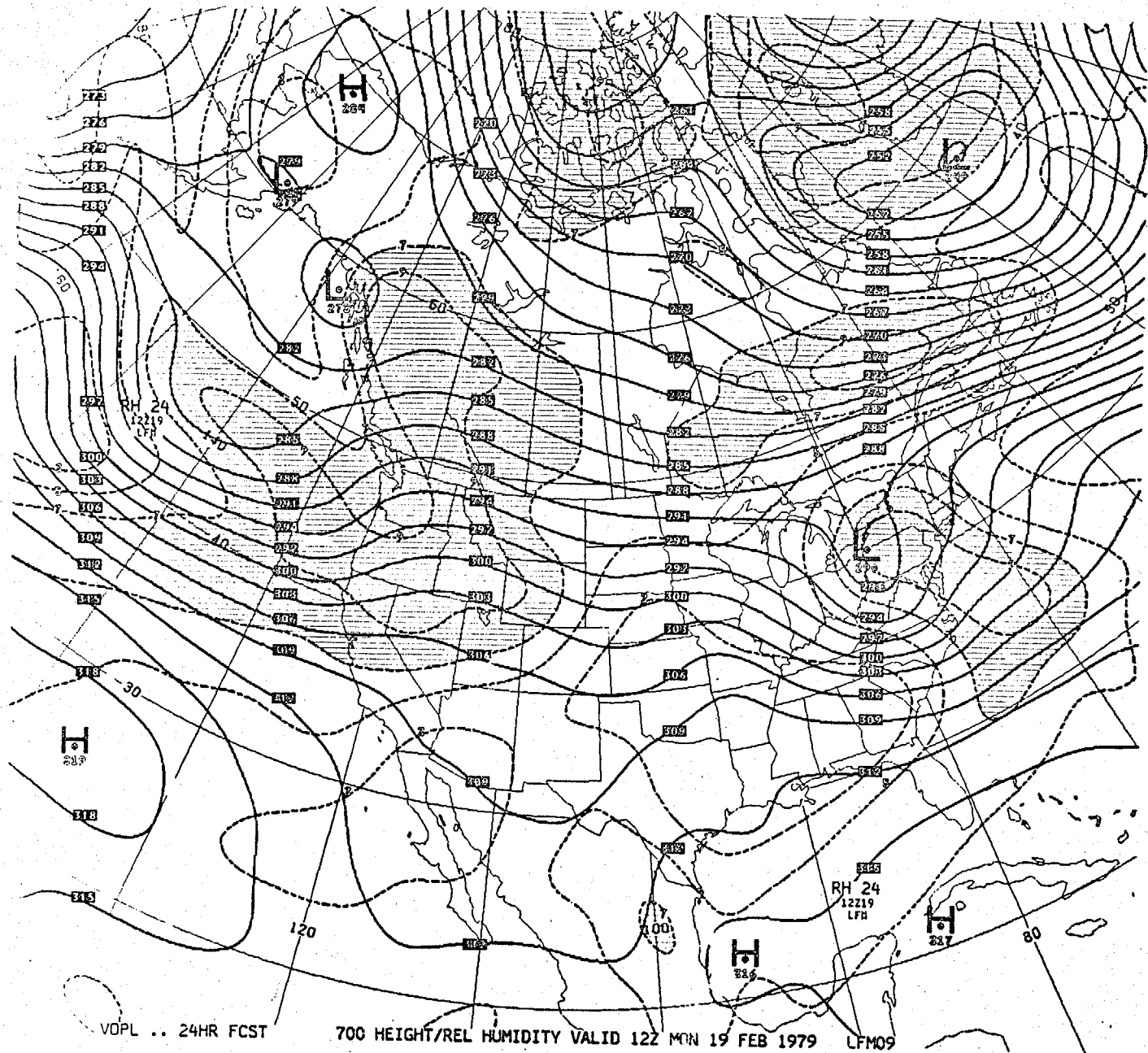
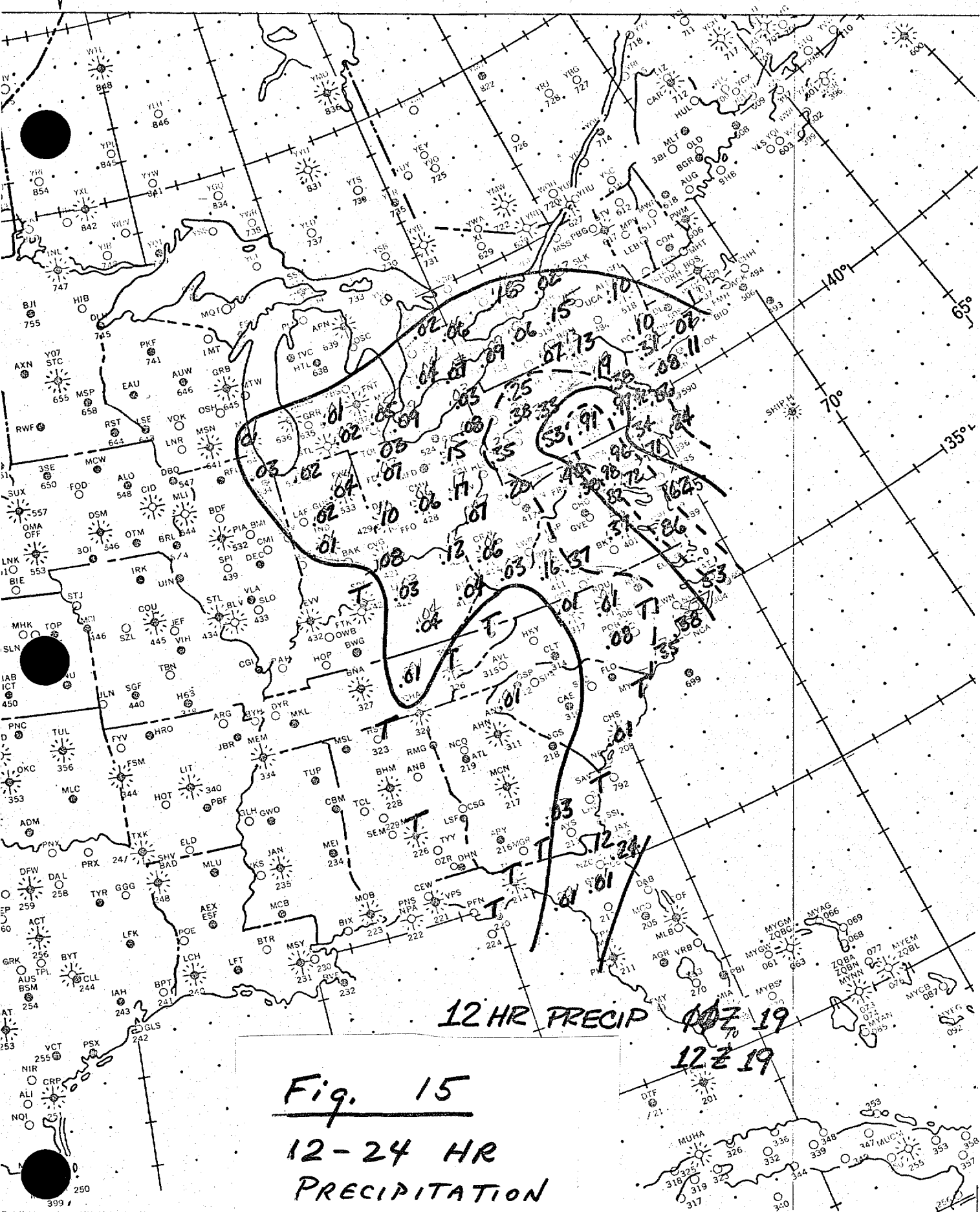
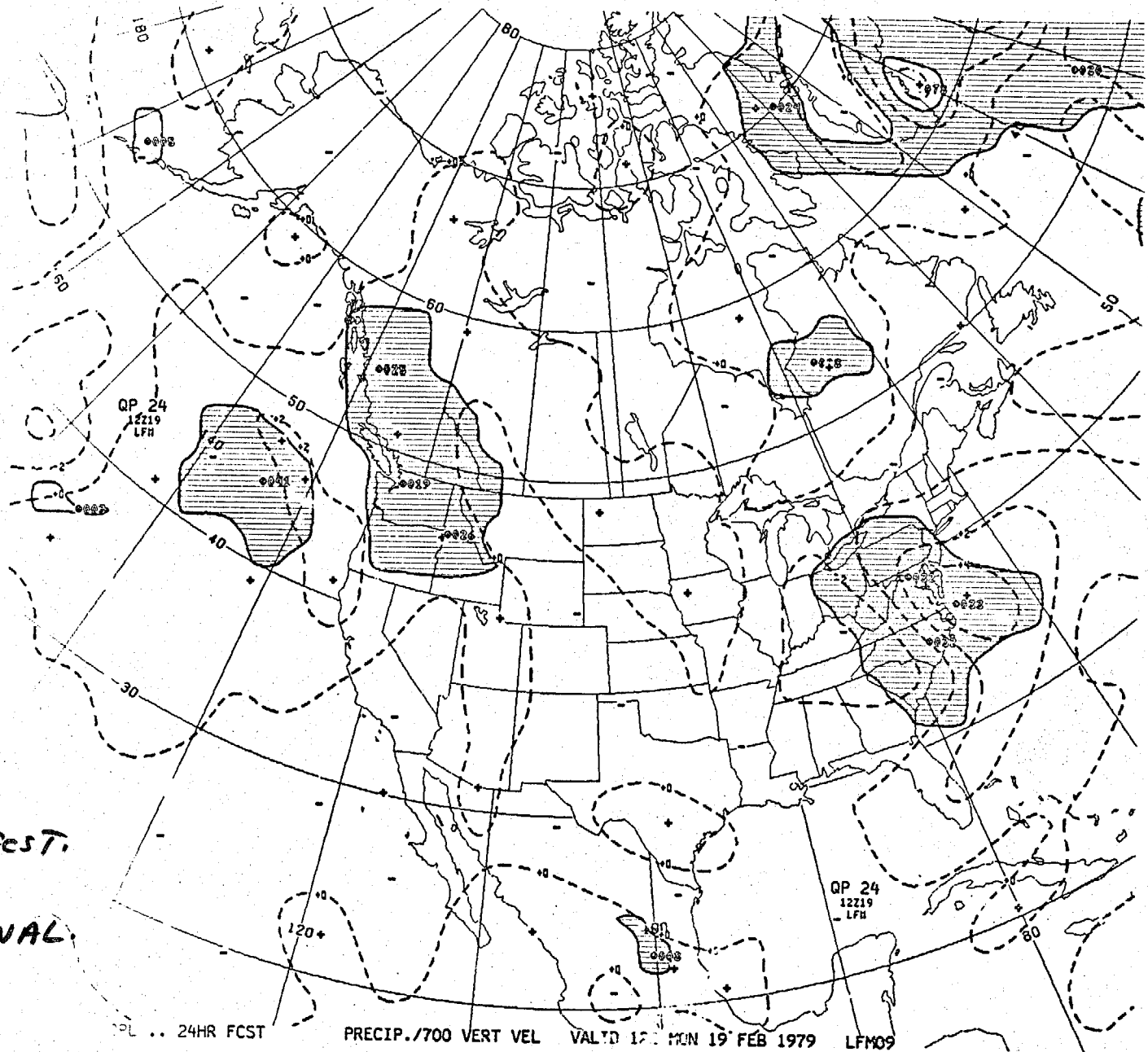
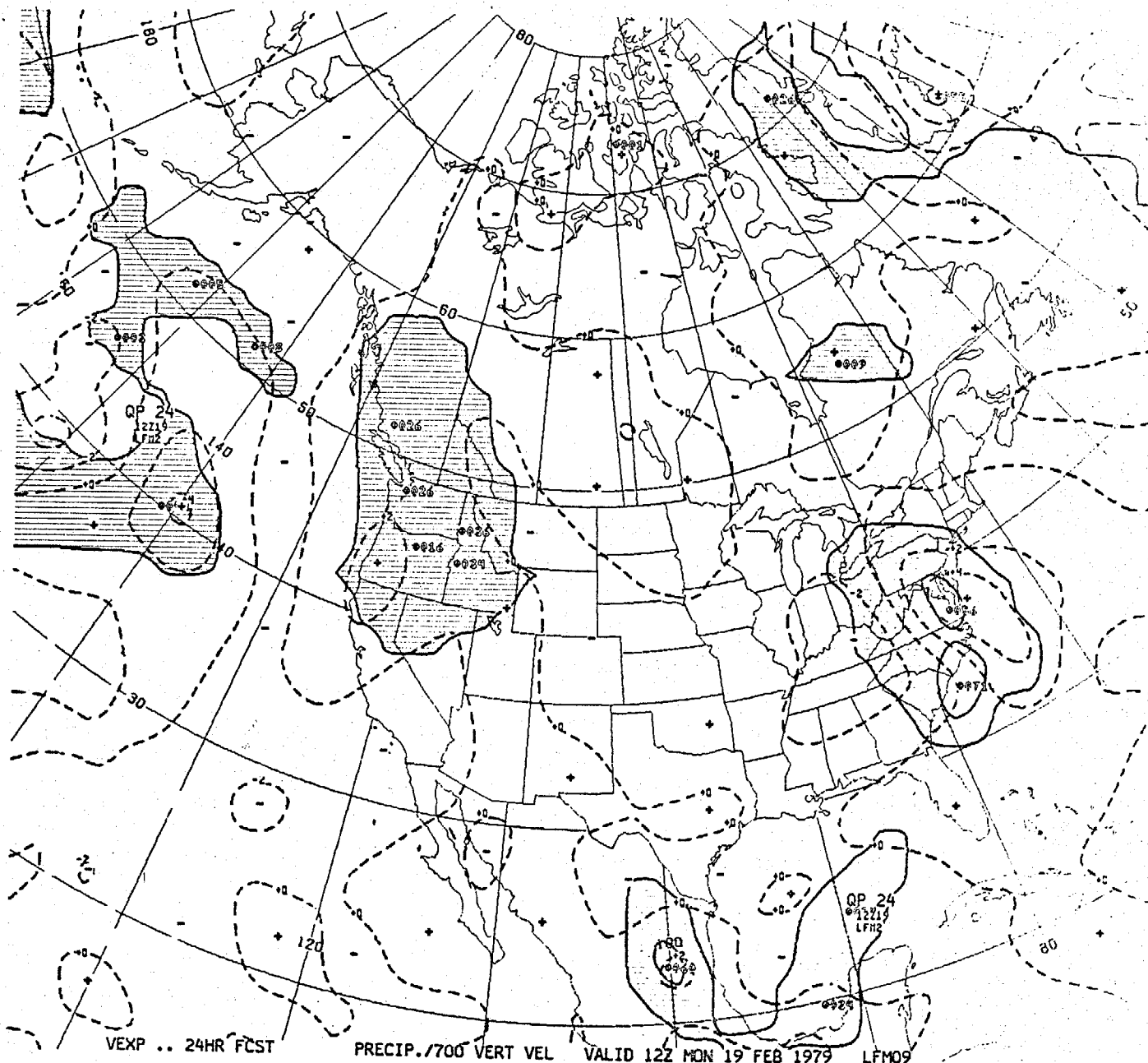


Fig. 13. LFM FCST. FROM OPNL. LFM ANAL.

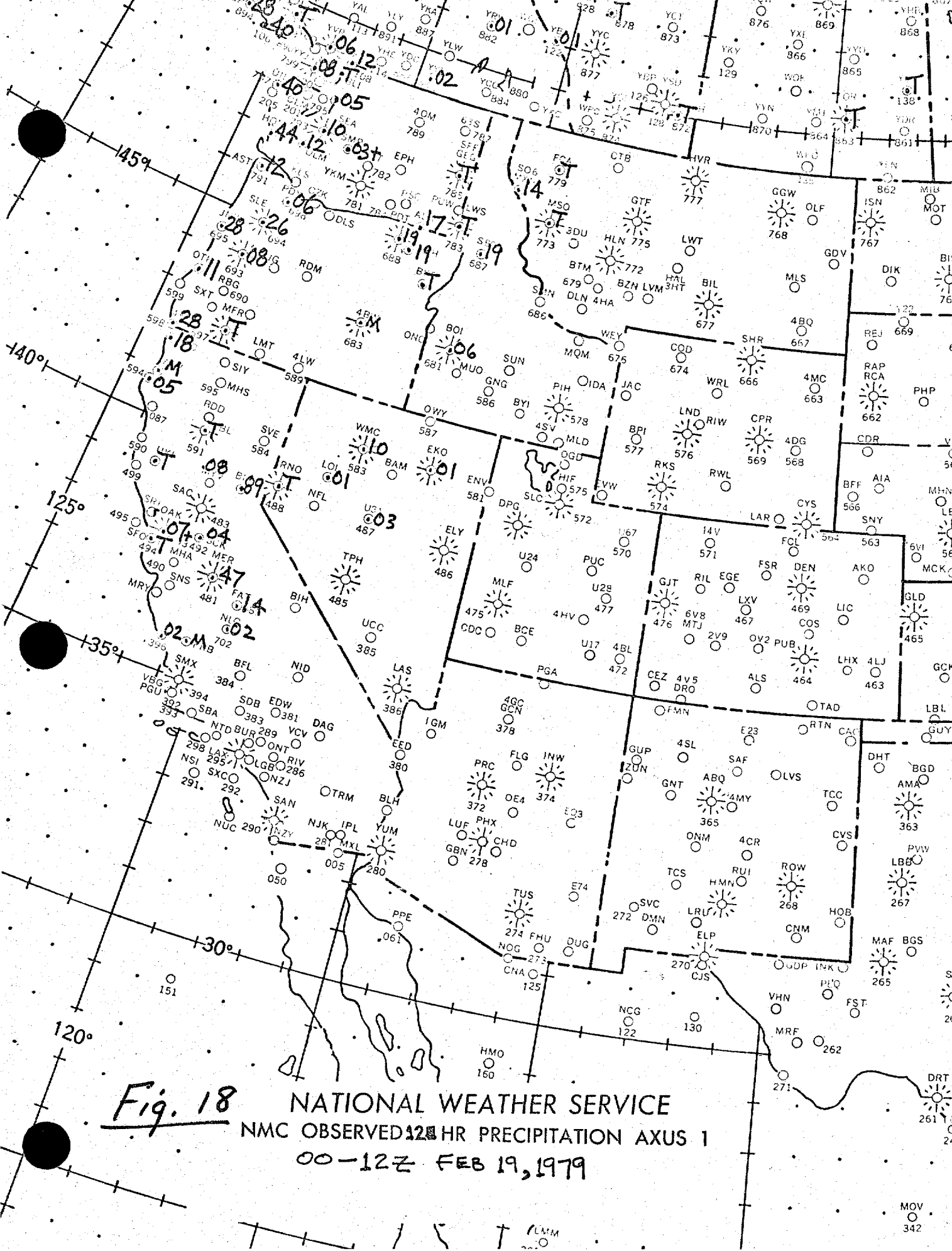


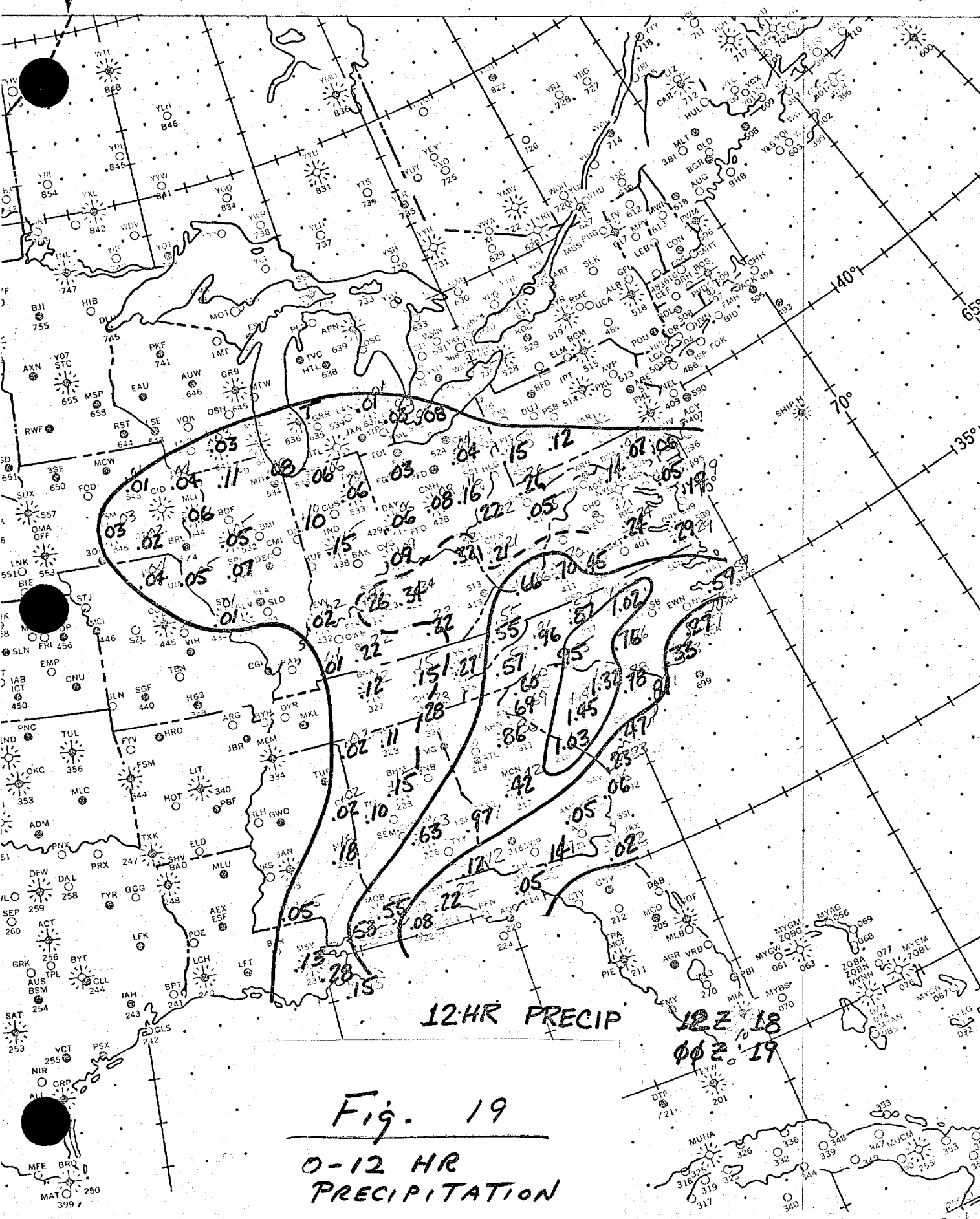
16





~~SI (F)~~ Fig. 17. LFM FCST. FROM MOI ANAL. (SMOOTHED)





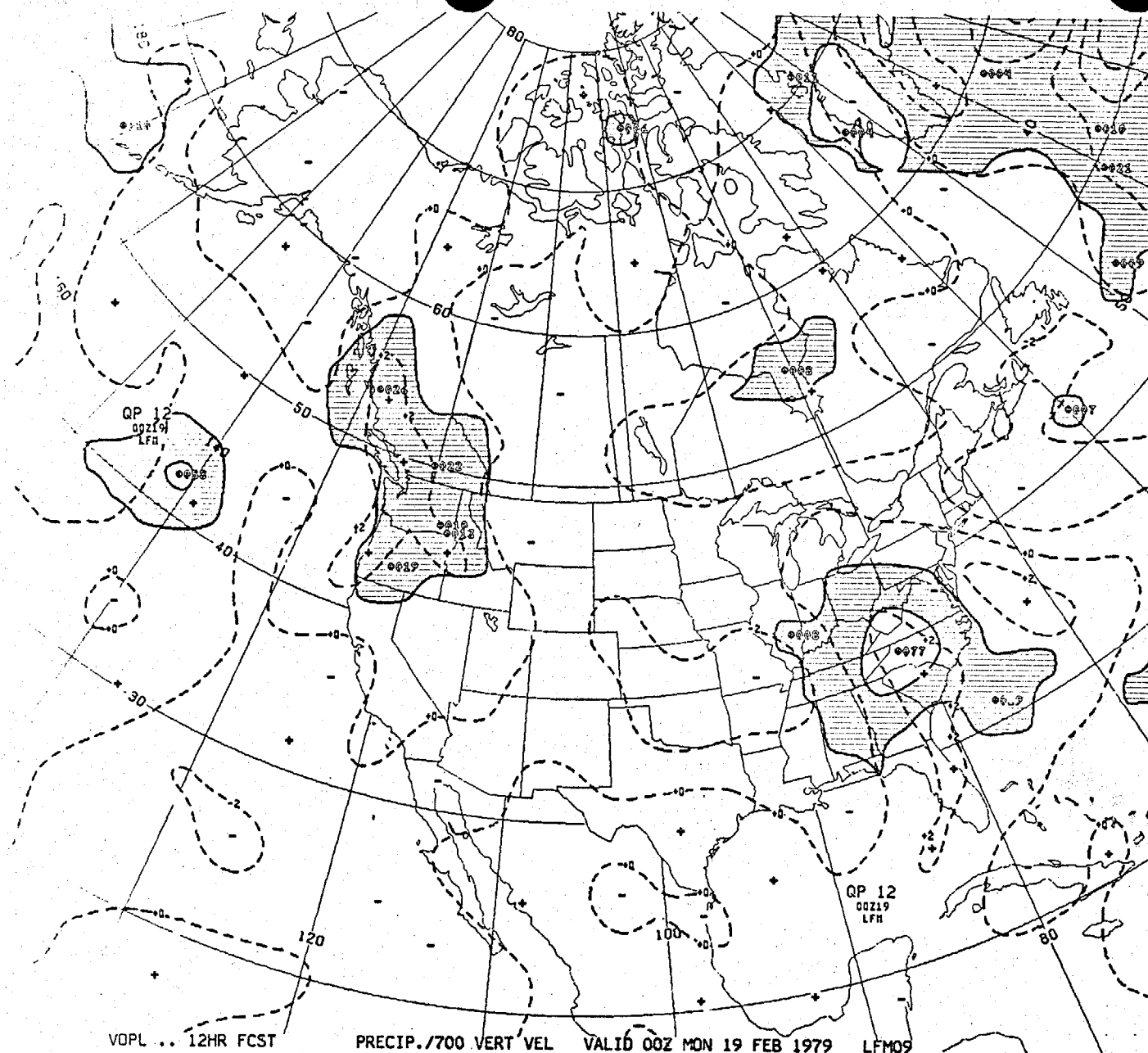
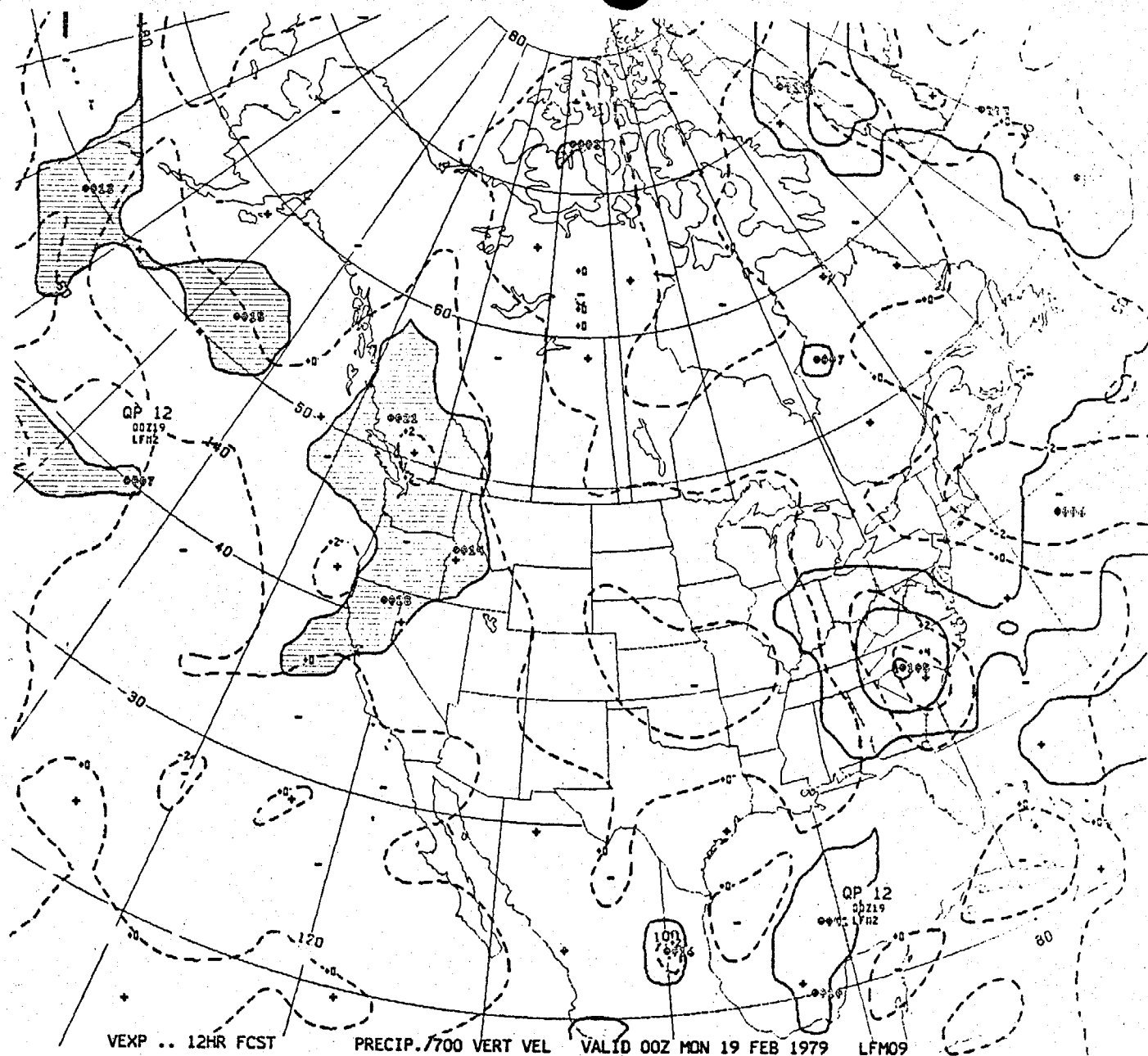
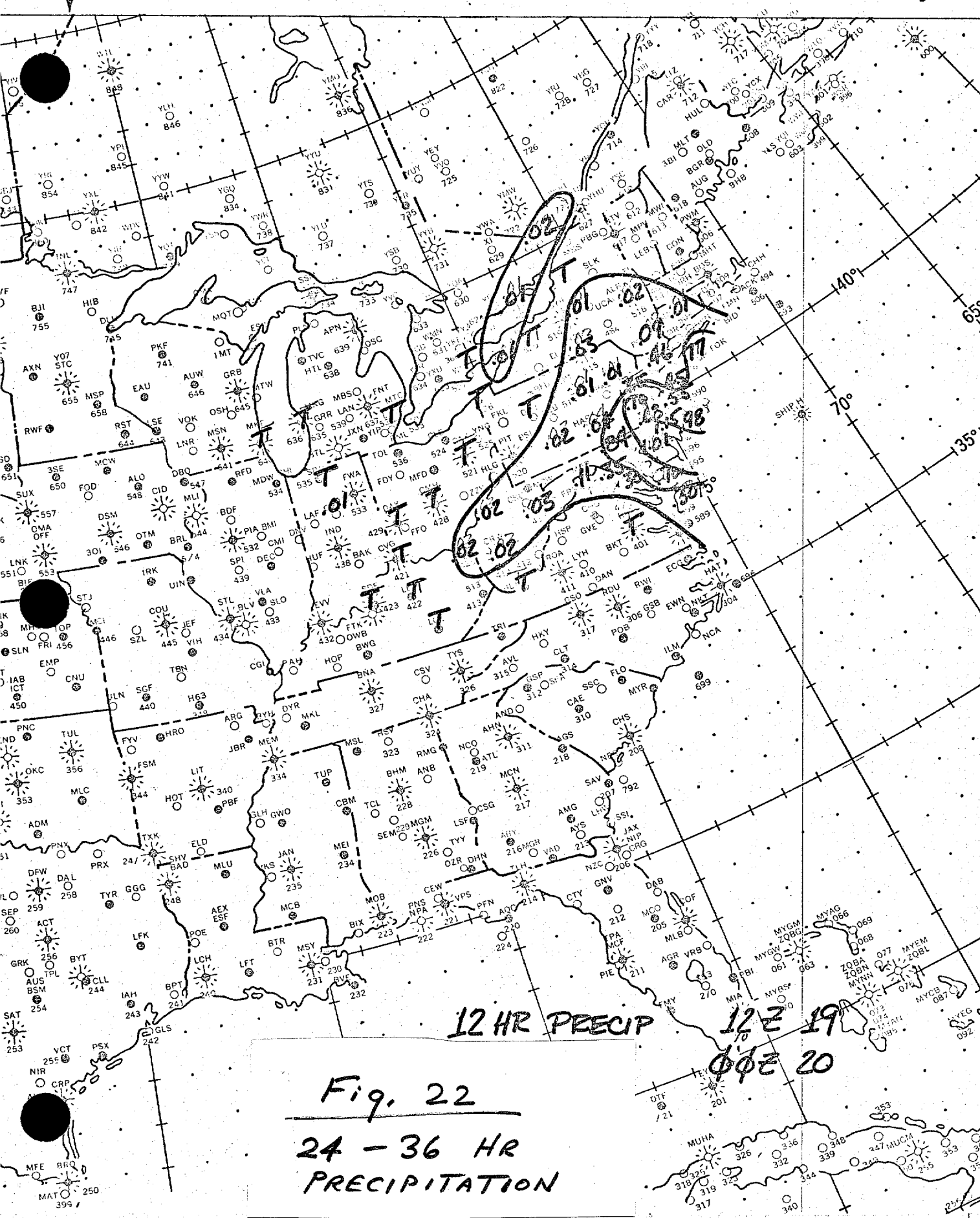


Fig. 20. LFM FCST. FROM OPNL, LFM ANAL.



~~DI (F)~~
 Fig. 21 LFM FCST. FROM MOI ANAL. (SMOOTHED)



23

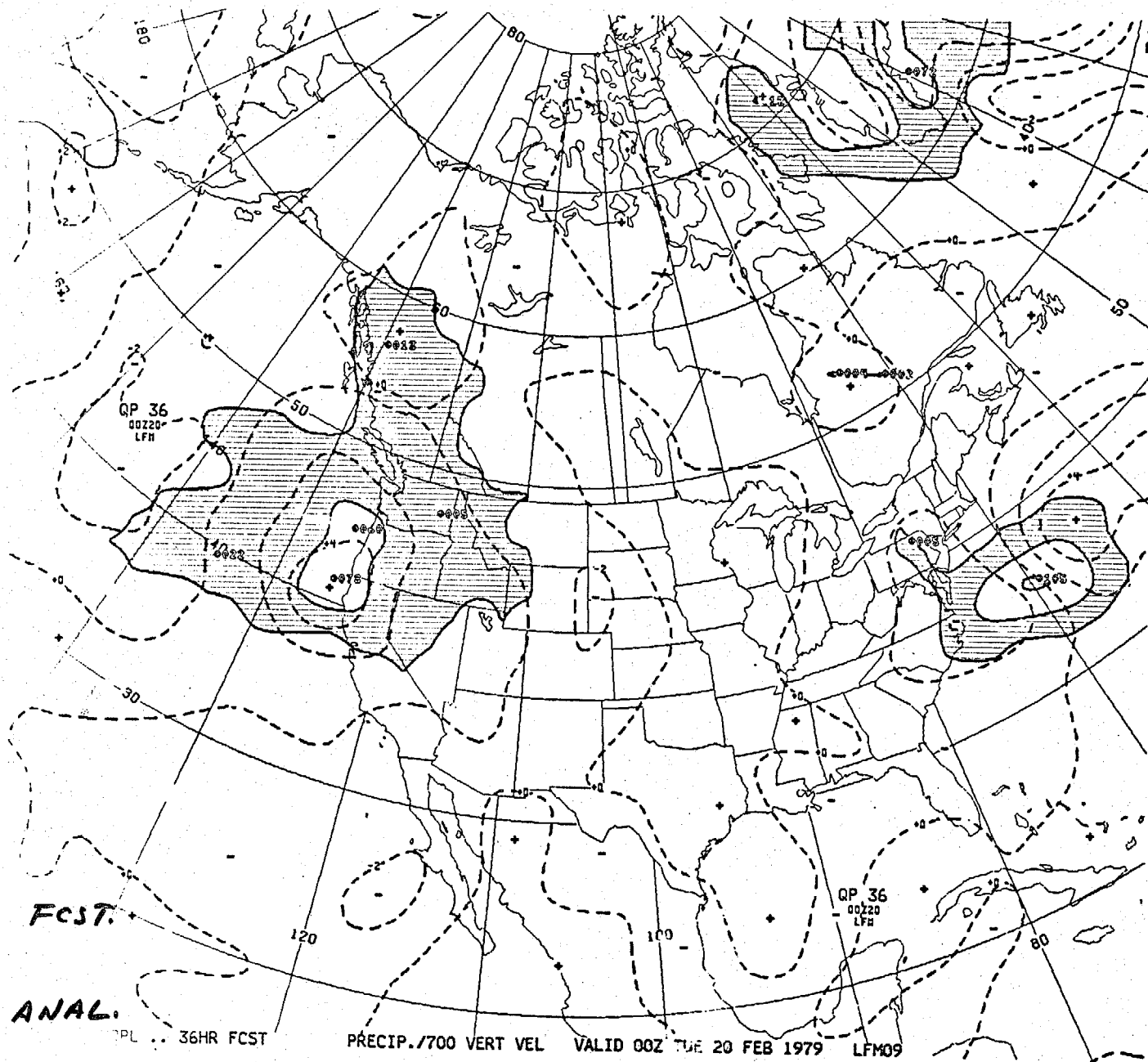


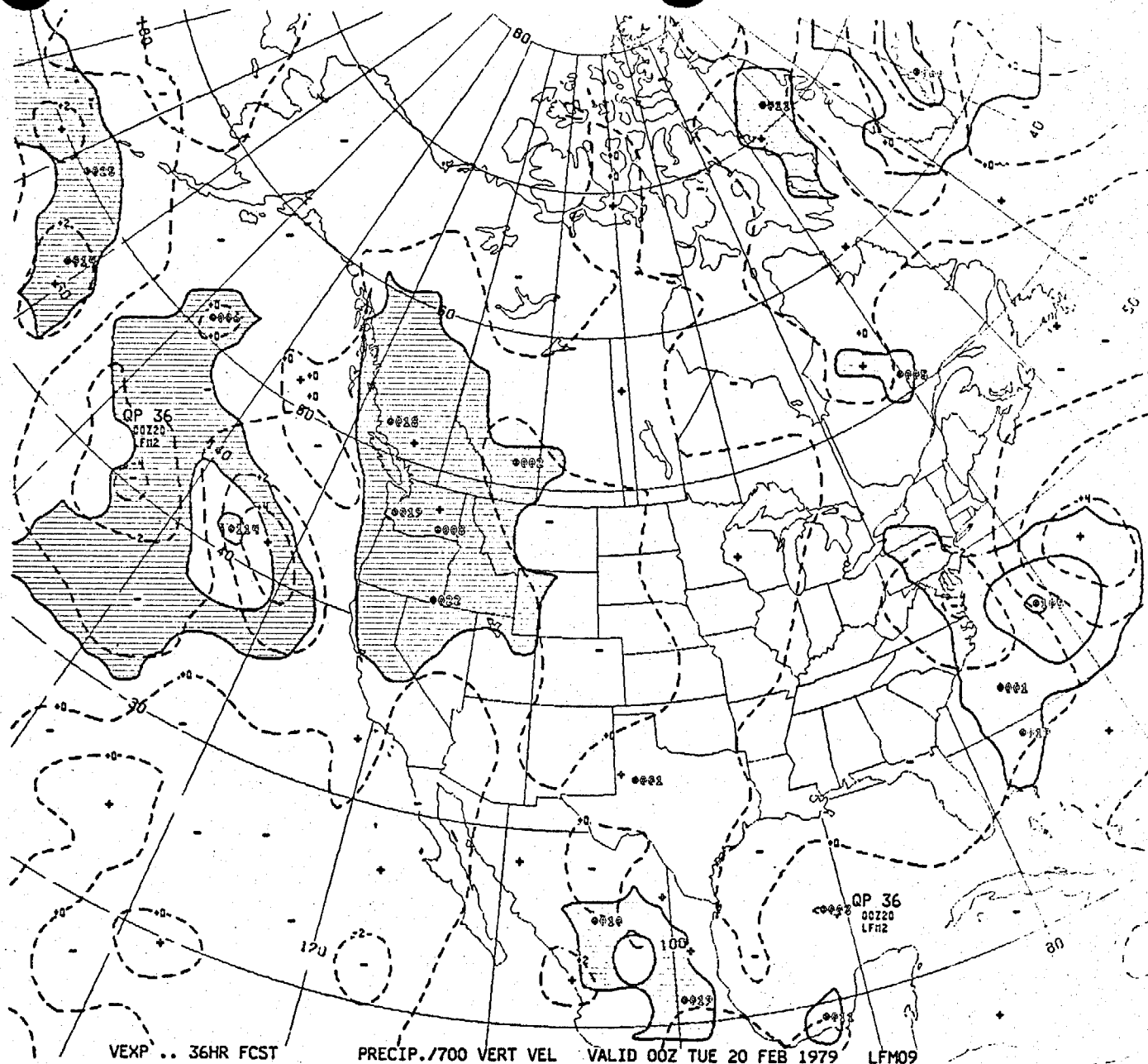
Fig. 23
OPNL. LFM FCST.
FROM
OPNL. LFM ANAL.

PL ... 36HR FCST

PRECIP./700 VERT VEL

VALID 00Z TUE 20 FEB 1979

LFM09



OT (F)

Fig. 24
OPNL. LFM
FORECAST
FROM
MOI ANALYSIS
(SMOOTHED)

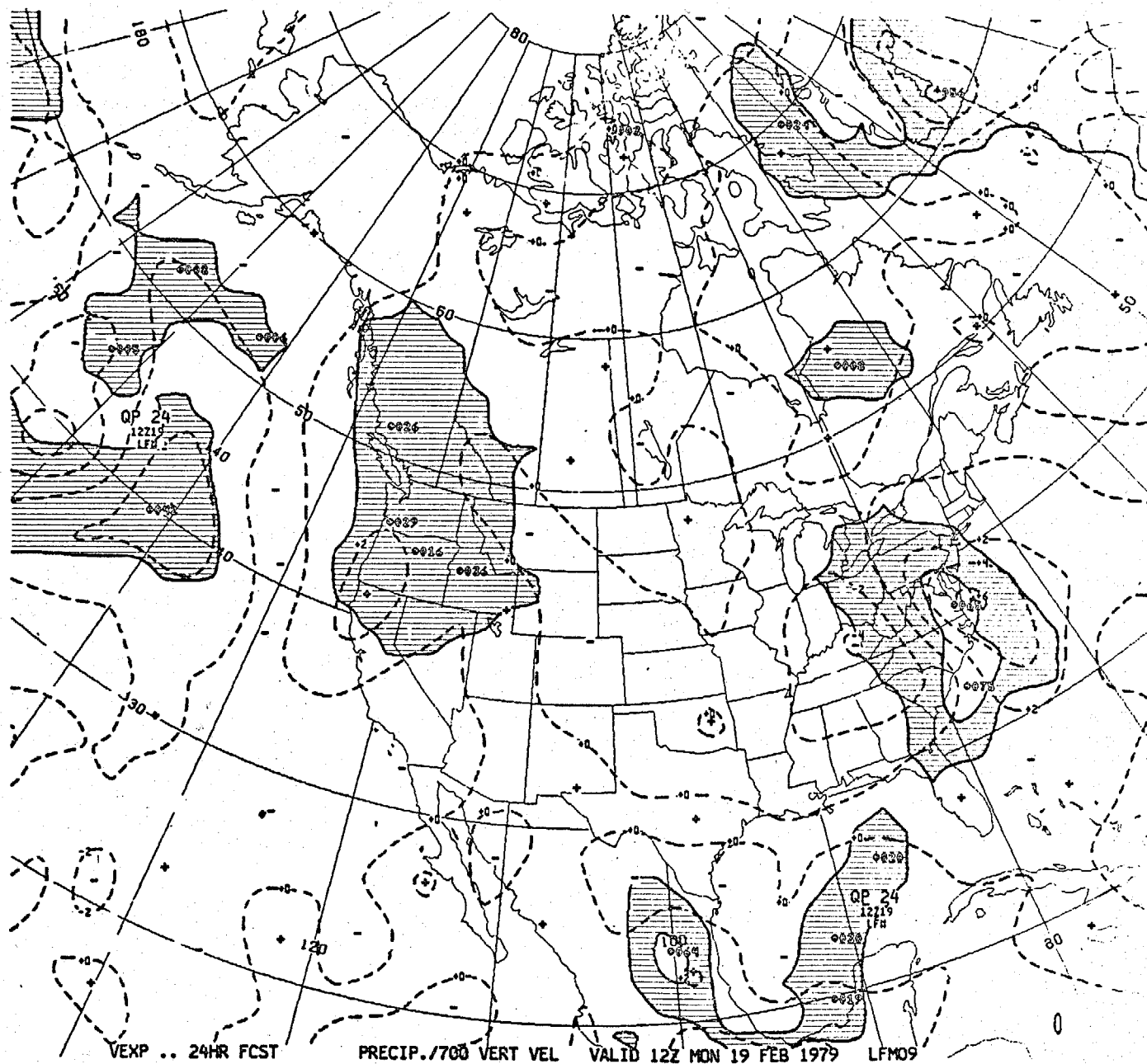
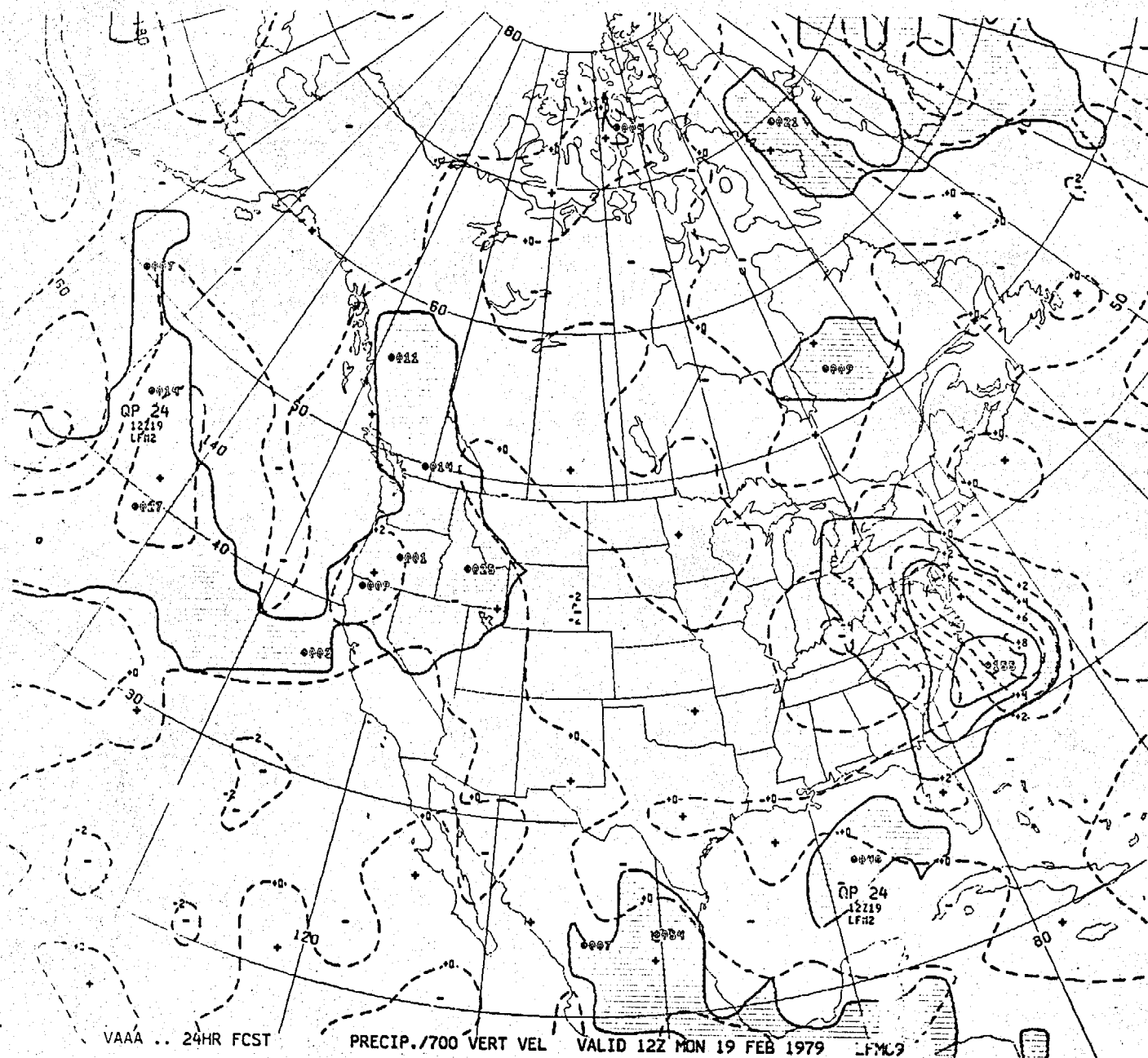


Fig. 25
OPNL LFM
FORECAST
FROM
MOI ANALYSIS
(UNSMOOTHED)

Fig 26

Fig. 26
DEAVEN'S LFM
MODEL VERSION
FROM
MOI ANALYSIS
(UNSMOOTHED)



VAAA ... 24HR FCST

PRECIP./700 VERT VEL

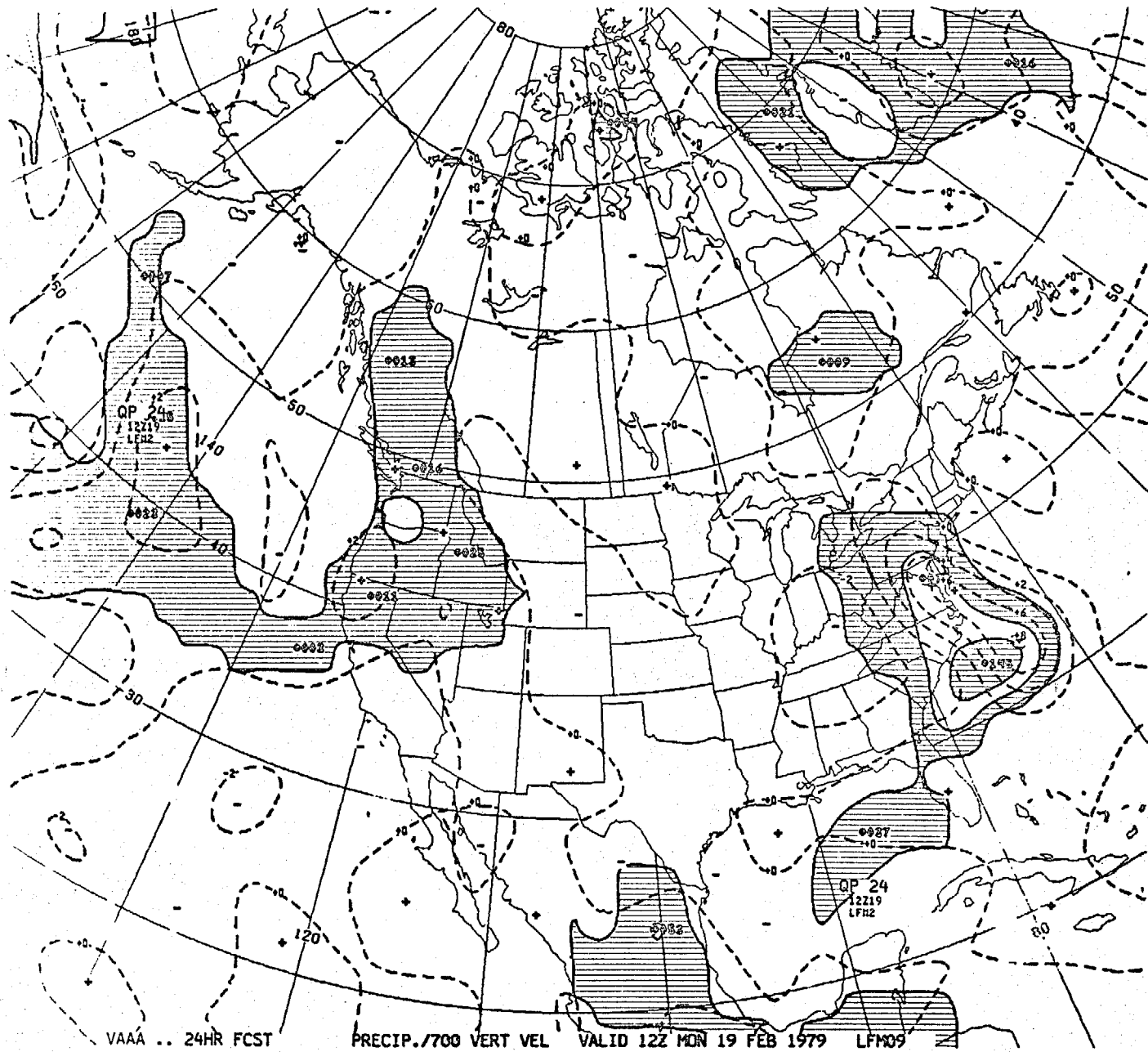
VALID 12Z MON 19 FEB 1979

LFM2

OT / D

Fig 27

Fig. 27
DEAVEN'S LFM
MODEL VERSION
FROM
MOI ANALYSIS
(SMOOTHED)



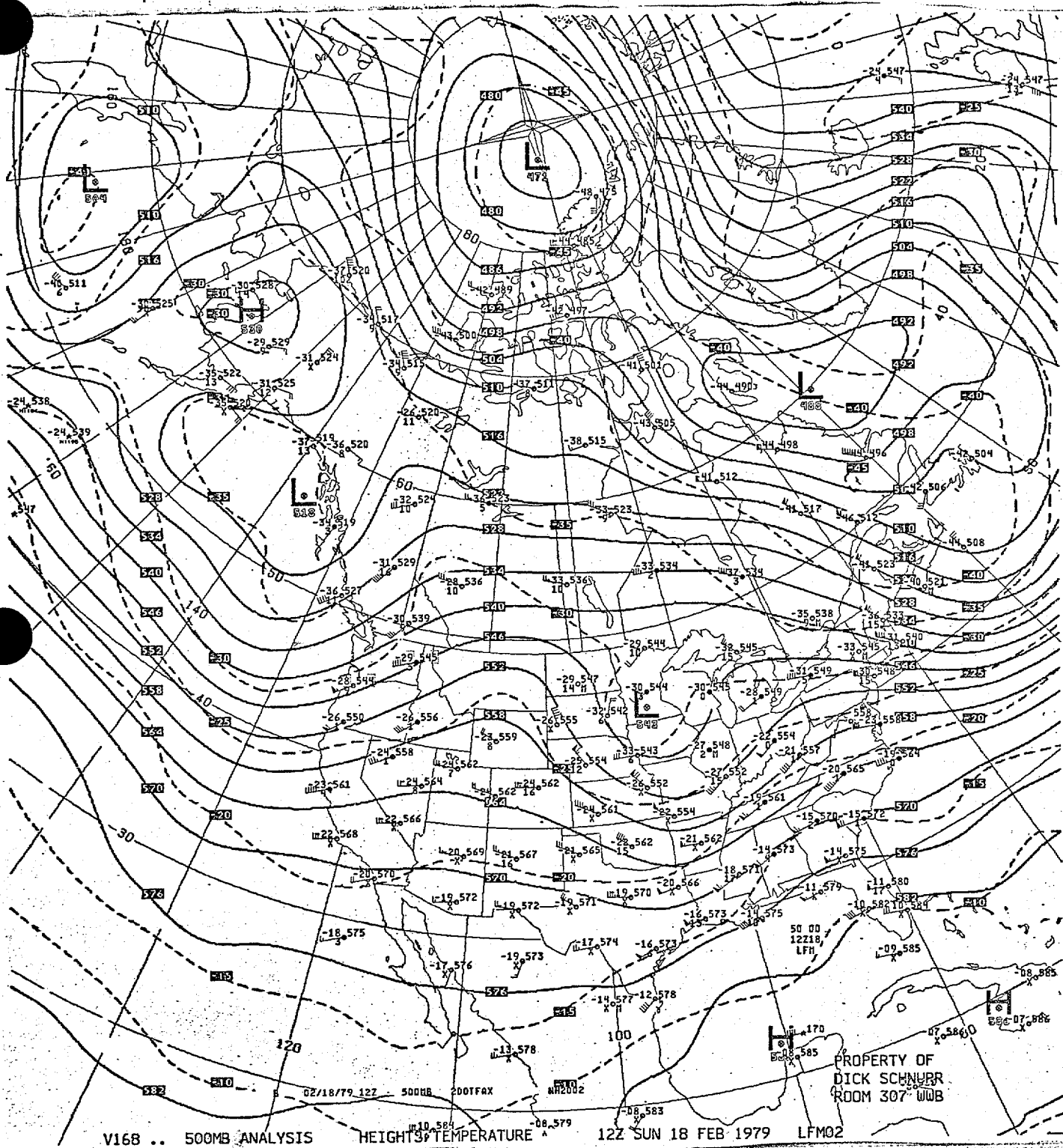


Fig. 4. SMOOTHED MOI ANALYSIS

The three-dimensional correlation function for RH is assumed to be the product of a two-dimensional lateral correlation function and a one-dimensional vertical correlation, thus

$$\rho_{ij}^{rr} = \mu_{ij}^{rr} v_{ij}^{rr} \quad (4.1)$$

The lateral correlation has the form

$$\mu_{ij}^{rr} = \exp\{-k_h[(\Delta\tau_{ij})^2/a + (\Delta\eta_{ij})^2]\} \quad (4.2)$$

where $\Delta\tau_{ij}$ is the component of separation distance between points i and j which is parallel to the vector wind at the grid point and sigma level of the updated wind field analysis and, $\Delta\eta_{ij}$ is the component of separation distance normal to this direction. (The vector wind at the two upper sigma levels is obtained by vertical interpolation of the vector wind from the two bracketing mandatory levels. For the lowest sigma layer, the forecast boundary-layer wind is used.) The value of k_h is the same as that used in the multivariate analysis, and the parameter a is defined by

$$a \equiv 1 + .004(u_g^2 + v_g^2); \quad a \leq 4; \quad (4.3)$$

where u_g and v_g are the components of the vector wind at the grid point and sigma level.

The above lateral correlation function has isopleths of equal magnitude which are elliptic, with major axis in the direction of the grid point wind. The eccentricity of the ellipse is determined by the value of a , which is not allowed to exceed 4. Thus in strong wind flow, the relative humidity is assumed to be correlated up to four times as strongly in the direction of wind flow than in the normal direction. In weak flow, the correlation is nearly isotropic. The number .004 appearing in (4.3) was determined empirically.

The vertical RH correlation is modeled by

$$v_{ij}^{rr} = 1 + (k_p + c) \ln^2(p_i/p_j)]^{-1}, \quad (4.4)$$

where

$$c \equiv -75 \left(\frac{\partial \theta}{\partial p} \right)_g; \quad \left(\frac{\partial \theta}{\partial p} \right)_g \leq 0 \quad (4.5)$$

Here $(\partial \theta / \partial p)_g$ is the static thermal stability for the grid point and sigma layer. It is determined from the updated temperatures at bracketing mandatory levels for the two upper layers, and by the 1000-850 mb temperature profile for the lowest layer. The number 75 is an empirical constant.

Correlation (4.4) has the effect of increasing the vertical depth of RH correlation in areas of low thermal stability and restricting it where the air mass is very stable. The reasoning behind this is that relative humidity profiles typically show more layering when the stability is greater.

The weights which the observations receive in the analysis are also a function of the normalized rms errors assigned to them. These are listed by instrument type in the last column of Table 1. Here it can be seen that the pseudo-RH values obtained from the cloud and weather reports at surface synoptic stations are assigned an error level twice that of rawinsonde RH values. Nevertheless, the RH analysis is primarily a function of the surface reports because there are so many more of these than there are rawinsonde soundings. A suggested experiment which was tried for one regional analysis (on the hemispheric grid) was the combining of all surface and surface bogus (from NESS) cloud/weather reports within a grid box into one "super-observation." Comparison of this experimental analysis with the regular method described above showed little difference in the moisture fields but did save considerable computer time.

5. Design Features of the Analysis Scheme

a. Data Base and Analysis Grid

The data base for the regional MOI analysis in order to initialize the LFM model consists of all data currently available at the "early" data cutoff (approximately 1 hour after observation time). This consists of mandatory level rawinsonde data, surface synoptic data, aircraft data, satellite cloud tracked winds, and (previously) satellite VTPR mandatory level temperatures. The codes are designed to accept other new kinds of data, such as the new TIROS-N data. The data can be assigned quality indicators; at present, such indicators are assigned only to rawinsonde data.

The 12-hr hemispheric numerical prediction which provides the guess fields for the analysis is interpolated to an extended version of the LFM-I grid by utility routine W3FT00 (J. McDonell). The analysis is then interpolated to the locations of the available observations. Linear interpolation is used at constant pressure, and linear interpolation in the logarithm of pressure is used for observations not at mandatory pressure levels. Residuals (prediced minus observed values) are then computed. Since the analysis at border grid points should depend on observations external to the LFM grid as well as within it, the grid is extended by a "moat" of 12 additional LFM-I grid points on each side. Residuals are also computed for observations located within this moat.

The 1000-mb heights and winds generated from surface reports are then used in the analysis of the 1000-mb height field (which analysis may be converted back into a sea-level pressure field) and in the height, temperature, and wind analyses at 850 mb, 700 mb, and (rarely) higher mandatory pressure levels. Additionally, upper air data, if selected by the data search and screening routine, may be used multivariately in the 1000-mb height analysis. The LFM initializer does not require 1000-mb temperature or wind analyses.

d. Data Search and Selection Procedure

The data search and selection procedure for the regional MOI scheme is very similar to that of the global scheme. A preanalysis program sorts the observed data by location on the extended LFM-I grid. Then a search is made of the data about each grid point to be updated. Searches are carried out successively in nested rectangles of 8, 16, and 24 grid lengths on a side centered on the grid point. At least one sounding, either rawinsonde or satellite, is required in each of three quadrants for the search to be satisfied. Otherwise the largest rectangle of data is used. In practice, the smallest or intermediate sized rectangles are used over about 75% of the grid, the largest rectangle for the remaining 25% of the grid.

Once a search rectangle has been selected, all data within the atmospheric volume, up to 100 mb, delineated by the rectangle are screened for possible use in the MOI analysis of each of the four variables at each mandatory level from 850 to 100 mb, and for the geopotential height at 1000-mb. As in the global scheme, the screening is based on the magnitude of the correlation, as modified by the estimated observational error, between the observation location and the grid point and pressure level to be analyzed. Cross correlations are multiplied by a factor of 1.75 for the screening only. The 10 observations which have the highest correlation magnitude thus determined are actually used in the analysis. (Some experimental analyses indicate that running time could be reduced without compromising analysis performance if the maximum number of observations used were reduced to 8.)

There are two restrictions currently imposed on the use of all the data in a fully multivariate form. Because rawinsonde temperatures and heights to a large degree provide redundant information, only the heights, if reported, are used in the height and wind analyses, and the temperatures are used only in the temperature analyses. But at a rawinsonde mandatory level where, say, the height is missing but a temperature is reported, the temperature may be used in the height and wind analyses.

The other restriction is that, for satellite temperature/ height soundings, only the temperatures are used in the analysis. The present scheme has no way to add 1000 mb analyzed heights to the reported heights since a 1000-mb analysis is not done separately, and before, the upper-air analyses.